

Horst Feldmann



# Der Griff nach den Genen

Von den Anfängen der Genetik  
zu den heutigen Genomprojekten

2014

# Inhaltsübersicht

## Vorwort

### 1 Wissenschaftliche Genetik und Molekulargenetik

- 1.1 Gene – Die Erbträger
- 1.2 Gene – Die Informationsträger
- 1.3 DNA – Der Stoff aus dem die Gene sind
- 1.4 Rekombinante DNA – Gene in Substanz
- 1.5 Reverse Genetik

### 2 Genomprojekte

- 2.1 Gene und Genome
- 2.2 Die Bäckerhefe als eukaryotisches Modellsystem
  - 2.2.1 Das Hefegenom wird sequenziert
  - 2.2.2 Analyse von Genfunktionen
  - 2.2.3 Im Hefegenom finden sich menschliche Krankheitsgene
  - 2.2.4 Hefen als Modellsysteme für die Evolution
- 2.3 Genome von anderen Mikroorganismen
- 2.4 Genome von weiteren, genetisch relevanten Modellsystemen
- 2.5 Das Humangenomprojekt
  - 2.5.1 Vorgeschichte und Konzept
  - 2.5.2 Strategien und Erfolge
  - 2.5.3 Neue Aspekte im Humangenom-Programm
- 2.6 Andere Genomprojekte im Fortschritt
  - 2.6.1 Menschliche Parasiten
  - 2.6.2 Nutzpflanzen
  - 2.6.3 Tiere
  - 2.6.4 Mikroben
  - 2.6.5 Begonnene und laufende Genomprojekte
    - 2.6.5.1 Künstliche Bakterienzellen
    - 2.6.5.2 'Synthetic Genomics'
    - 2.6.5.3 Human-Mikrobiom-Projekt
    - 2.6.5.4 Erforschung aller Kleinlebewesen der Weltmeere
    - 2.6.5.5 Das erste diploide Genom eines einzelnen Individuums
    - 2.6.5.6 Südafrika Genomprojekt

### 3 Gene als Material

- 3.1 Manipulierte Gene – Chancen und Risiken
  - 3.1.1 Biotechnisch erzeugte Genprodukte
  - 3.1.2 Anwendungen der Biotechnologie in der Medizin
- 3.2 Transgene Organismen
  - 3.2.1 Transgene Tiere
  - 3.2.2 Gentherapie beim Menschen
  - 3.2.3 Transgene Pflanzen
- 3.3 Weiße Biotechnik
  - 3.3.1 Hefen als Bioreaktoren
  - 3.3.2 Bakterien als Bioreaktoren

### 4 Genfunktionen, Gengeflechte und Netzwerke

- 4.1 Genomics – Die gesamten Gene und ihre Funktionen in einem Organismus
  - 4.1.1 'Genomics' mithilfe der Mikrochip-Technik
  - 4.1.2 Andere Sparten von „-omics“
- 4.2 Wechselbeziehungen zwischen Genen und Proteinen
- 4.3 Nicht-codierende Gene im Genom
  - 4.3.1 RNA-Interferenz
  - 4.3.2 Andere nicht-codierende RNAs
- 4.4 Regulation der Zellaktivität

## Ausblick

## Literatur

## Glossar

# Der Griff nach den Genen

## *Von den Anfängen der Genetik zu den heutigen Genomprojekten*

**Horst Feldmann**

*Adolf-Butenandt-Institut der Ludwig-Maximilians-Universität München*

### Vorwort

Mein erster Aufsatz zu diesem Thema datiert aus dem Jahre 1999 und entstand anlässlich eines Vortrages an der Universität Kassel. Er wurde wiederholt überarbeitet, um jeweils aktuelle Daten einzufügen. Der Artikel versucht nachzuzeichnen, wie sich nach der Entdeckung von Gesetzmäßigkeiten in der Vererbung durch Gregor Mendel die Wissenschaftliche Genetik seit etwa Mitte des 19. Jahrhunderts weiter entwickelt und seit Mitte des vorigen Jahrhunderts in immer schnelleren Schritten zu revolutionierenden Entdeckungen, Techniken und Einsichten geführt hat. Heute erlauben uns diese Erkenntnisse, individuelle Gene in Substanz zu isolieren und zu manipulieren, so dass wir bereits beträchtliche Einblicke in die Erbinformationen zahlreicher Organismen auf molekularer Ebene besitzen. Unausweichlich waren in dieser Entwicklung Abwägungen über den Nutzen gegenüber den Risiken, die der erreichte Kenntnisstand und die jetzigen Möglichkeiten der Gentechnik (oder Biotechnik) mit sich bringen. Sie sollen berührt werden, aber nicht im Vordergrund stehen. Vielmehr ist beabsichtigt, wichtige Fakten herauszustellen, die zum jetzigen Stand der Entwicklung geführt haben, und deren Kenntnis eine Basis in allen Debatten in diesem Bereich bieten könnte.

## 1 Wissenschaftliche Genetik und Molekulargenetik

### 1.1 Gene - Die Erbträger

Von der Biologie her betrachtet, herrschte um die Mitte des 19. Jahrhunderts noch ‚finsternes Mittelalter‘. Selbst als *Louis Pasteur* nachgewiesen hatte, dass auch Mikroben Lebewesen sind, behaupteten viele seiner Zeitgenossen immer noch, Bakterien und Pilze seien Nebenprodukte chemischer Prozesse der Fäulnis.

Auch *Gregor Mendels* berühmte Kreuzungsversuche an Erbsen, die er 1865 beschrieb (sie brachten ihm den Spitznamen „Erbsenzähler“ ein; Abbildung1), wurden erst zu Beginn unseres Jahrhunderts wieder entdeckt, richtig verstanden und nicht mehr als bloße Zahlenmystik abgetan. Mit der gedanklichen Aufteilung der Erbinformation in einzelne Einheiten und ihrer ‚freien Kombinierbarkeit‘ gilt *Mendel* als Begründer der Wissenschaftlichen Genetik. Danach beinhaltet das Erbgut, die genetische Information eines Organismus, eine Vielzahl von einzelnen Erbfaktoren, jeder Erbfaktor

steuert die Ausbildung eines bestimmten Merkmals und stellt damit zugleich die Einheit einer Funktion dar.



**Gregor Mendels Kreuzungs-experimente**





Samen		Blüte	Schote		Stamm	
Form	Keimblatt	Farbe	Form	Farbe	Anordnung	Größe
						
Grau rund	Gelb	Weiß	Voll	Gelb	Blüten verteilt Schoten axial	Lang
						
Weiß kraus	Grün	Violett	Knotig	Grün	Blüten axial Schoten verteilt	Kurz
1	2	3	4	5	6	7

Abbildung1: Mendel's Experimente von 1865. Oben: Mendel und sein Garten im Kloster Brunn. Unten links: Kreuzungsschema von Erbsen mit zwei dominanten Erbmerkmalen A=rund und B=gelb, und zwei rezessiven Erbmerkmalen a=grün und b=rau. In der ersten Kreuzung (F1 Generation) kommen ausschließlich die dominanten Merkmale zum Vorschein. Werden hieraus zwei Individuen gekreuzt, so entsteht das im Schema gezeigte Muster (F2 Generation). Unten rechts: Von Mendel untersuchte Vererbung von sieben verschiedenen Merkmalen von Erbsen.

Um 1900 wurden schließlich die Begriffe ‚Gen‘ für einen Erbfaktor und ‚Genom‘ für die Gesamtheit aller Erbfaktoren gebräuchlich. Die Veränderungen an Genen, die sich phänotypisch (d.i. als augenfällige Merkmale) zu erkennen geben, wurden als Mutationen bezeichnet.

## 1.2 Gene - Die Informationsträger

In der Zeit zwischen 1910 und 1915 konnte *Thomas Hunt Morgan* mit seiner Gruppe durch Arbeiten an der Taufliege *Drosophila* zeigen, dass Gene sich einzelnen ihrer Chromosomen zuordnen lassen und in linearer Weise auf den Chromosomen angeordnet sind (Abbildung 2). Schon in den ersten Experimenten konnte auf dem X-Chromosom dieser Fliege ein mutiertes Gen lokalisiert werden, das zu weißen statt zu normalen roten Augen führt. In den folgenden Jahren wurden viele weitere, an

einer Mutation erkennbare Gene lokalisiert, wobei sich die ‚**Genkartierung**‘ auf natürlich entstandene Genmutationen beschränken musste.



## ***Drosophila melanogaster*** und ihre Riesenchromosomen aus der Speicheldrüse

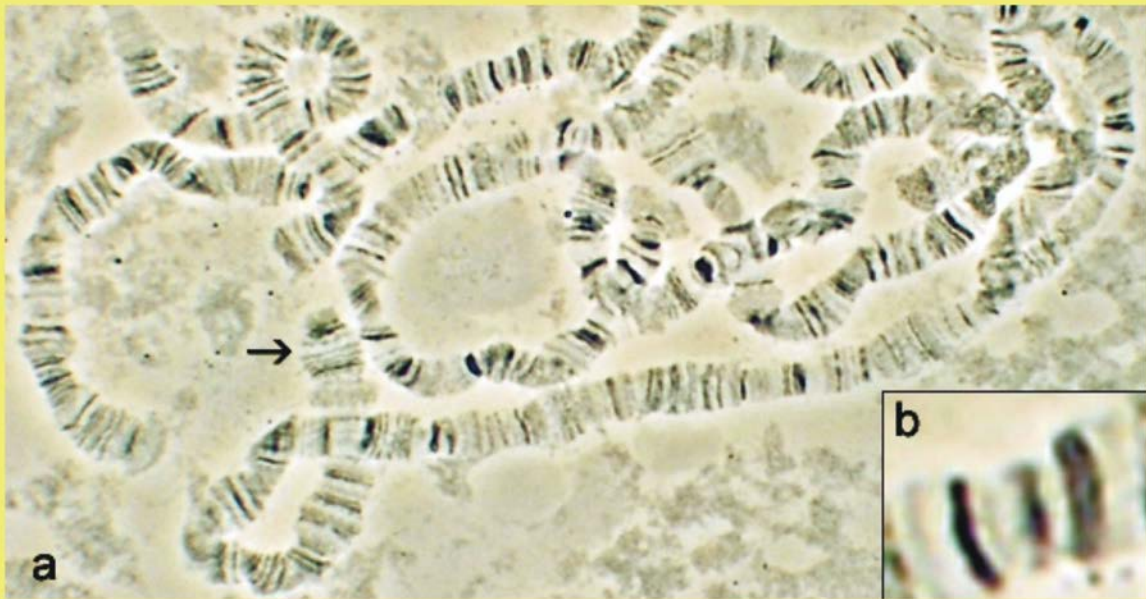


Abbildung 2: Die Taufliege *Drosophila melanogaster* und eine mikroskopische Aufnahme ihrer polytären Chromosomen („Riesenchromosomen“) aus der Speicheldrüse. Aus der Bänderung ließ sich eine detaillierte Genkarte erstellen.

In diese Zeit fällt auch die Entdeckung des englischen Arztes *Sir Archibald Garrod*, dass einige menschliche Erbkrankheiten auf Stoffwechseldefekte zurückzuführen sind, bei denen bekannte biochemische Reaktionen ausfallen. Er stellte die Hypothese auf, dass eine angeborene Stoffwechselanomalie auf das Fehlen eines spezifischen Enzyms zurückgehe, das von einem mutierten Gen nicht mehr hergestellt werden kann. War dieser Verdacht richtig, so musste für jedes Enzym, sogar für jedes andere Zellprotein, ein entsprechendes Gen vorhanden sein. Damals war über den Zellstoffwechsel jedoch wenig bekannt, und es dauerte noch etwa 30 Jahre, bis 1945, ehe *George Beadle* und *Edward Tatum* die ‚**Ein-Gen-ein-Enzym-Hypothese**‘ experimentell untermauern konnten. Sie hatten als Versuchsobjekt den Schimmelpilz *Neurospora crassa* gewählt, der auf einfachen Nährböden (aus Glucose und Mineralien) wächst, weil er praktisch alle anderen benötigten Nährstoffe selbst herstellen kann. Sie nutzten außerdem die seit 1926 bekannte Tatsache, dass Mutationen durch Röntgen- oder UV-Strahlung verstärkt induziert werden. Unter den isolierten Mutanten von *Neurospora* fanden sich zum Beispiel viele, deren Wachstum vom Zusatz jeweils eines



bestimmten Nährstoffes (z.B. einer Aminosäure, eines Vitamins, o. dgl.) zum Nährboden abhängig wurde. Jedes Mal stellte sich heraus, dass einer solchen Mutante eines der Enzyme fehlte, welche normalerweise für die zelleigene Synthese des betreffenden Nährstoffes benötigt werden.

### 1.3 DNA - Der Stoff aus dem die Gene sind

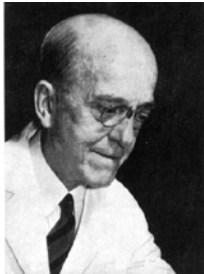


Abbildung 3:  
Oswald Avery 1937

Lange Zeit blieben 'Gen' oder 'Mutation' dennoch mehr oder weniger abstrakte Begriffe. Erst die Entdeckung von *Oswald Avery* (Abbildung 3), dass das genetische Material zur eigenständigen Vermehrung befähigter Organismen aus DNA (Desoxyribonucleinsäure) besteht, stellte eine stoffliche Basis her und eröffnete damit 1944 die Phase der Molekularen Genetik. Nucleinsäuren hatte *Friedrich Miescher* (Basel und Tübingen) bereits 1869 isoliert, deren Analyse später ergab, dass die DNA aus vier Basen (kurz als A,T,C und G bezeichnet) als Bausteinen besteht, die über ein sogenanntes Ribose-Phosphat-Rückgrat miteinander linear verknüpft sind. Dass also nach beinahe 80 Jahren sich das eher für eine Strukturkomponente gehaltene Molekül als informationstragendes Prinzip entpuppte, überraschte sogar Avery. Die gängige Meinung war damals nämlich, nur Proteine mit ihrer komplexen, dreidimensionalen Struktur könnten Erbinformation beinhalten.

### Struktur der DNA - Die Doppelhelix



James D. Watson



Francis H.C. Crick



Abbildung 4: Watson und Crick entdecken 1953 die „Doppelhelix“.

Die entscheidende Vorstellung über die Struktur der DNA lieferten dann *Wilkins, Watson und Crick* 1953. Die „**Doppelhelix**“, zwei über die komplementären Basenpaare A und T bzw. G und C miteinander verbundene DNA-Stränge (Abbildung 4), gestattete nun unmittelbar, die Mechanismen zur Weitergabe der genetischen Information von den Eltern auf die Nachkommen und gleichzeitig die Umsetzung dieser Information in der Zelle zu verstehen: Bei einer Zellteilung wird jeder DNA-Strang verdoppelt, es entstehen identische Kopien der genomischen DNA, so dass beide Zellen die gleiche Information erhalten. Diesen Vorgang nennt man auch **DNA-Replikation**; er ist heutzutage weitgehend molekular verstanden. Jeder der beiden DNA-Stränge wird möglichst genau mit einer Kopie versehen, so dass die Replikation ‚semikonservativ‘ verläuft, weil jeweils ein elterlicher Strang in der Kopie erhalten bleibt. Nach vollständiger Replikation können dann identische Kopien des Genoms auf die aus der Mutterzelle bei der Zellteilung entstehenden Tochterzellen verteilt werden.

Die Kurzformel für die weitere Umsetzung der genetischen Information wurde von Crick das '**Zentrale Dogma**' genannt (Abbildung 5): Es beinhaltet die Aussage, dass Information nur von den Genen über ihre Umschreibung in eine ‚Zwischenstufe‘ in Proteine umgesetzt werden kann. (Eine Ausnahme bilden RNA-Viren, die RNA als genetisches Material enthalten und in einer befallenen Zelle in DNA zurückgeschrieben werden müssen, um wirksam zu werden.)

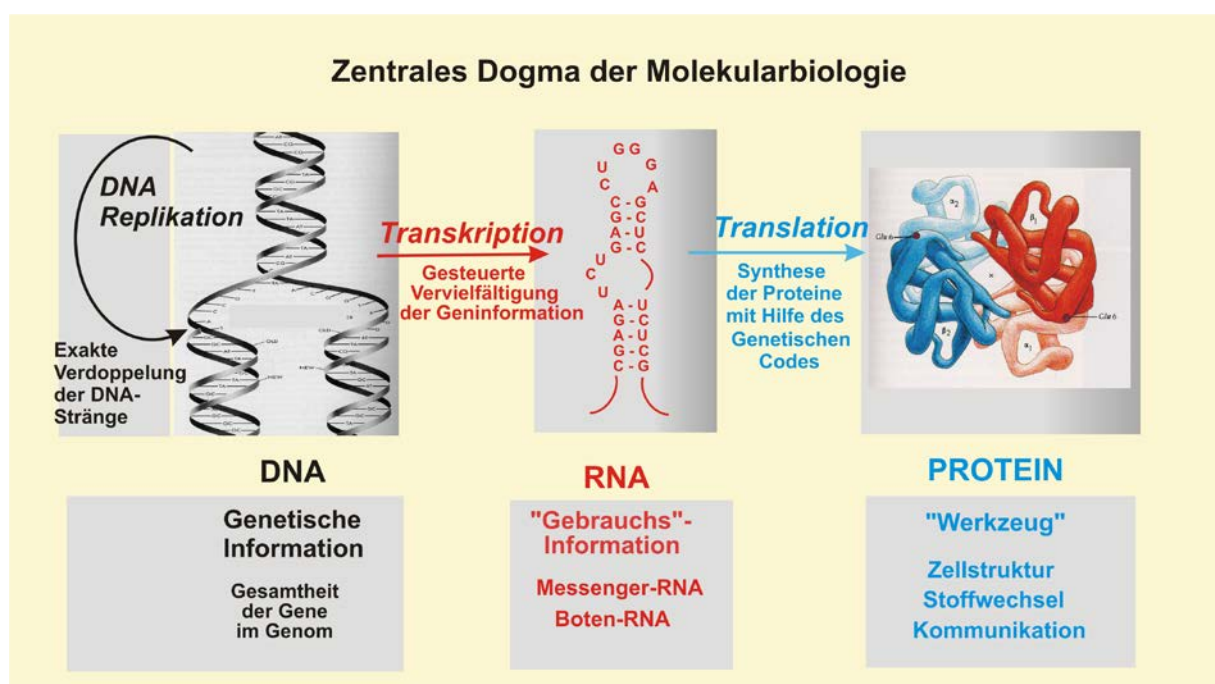
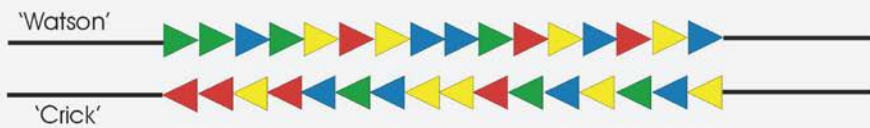


Abbildung 5: Zentrales Dogma der Molekularbiologie nach Crick. Die Weitergabe der genetischen Information kann nur ‚vorwärts‘, von DNA über RNA-Kopien in Proteine, erfolgen. Bei der Verdoppelung der DNA entstehen zwei identische Kopien; diese werden bei einer Zellteilung an beide Tochterzellen weitergegeben. Genetische Information wird bei der Herstellung von Proteinen dadurch abgerufen, dass Gene in (vielen) Kopien in RNA umgeschrieben werden, welche dann von der Proteinsynthese-Maschinerie benutzt werden.

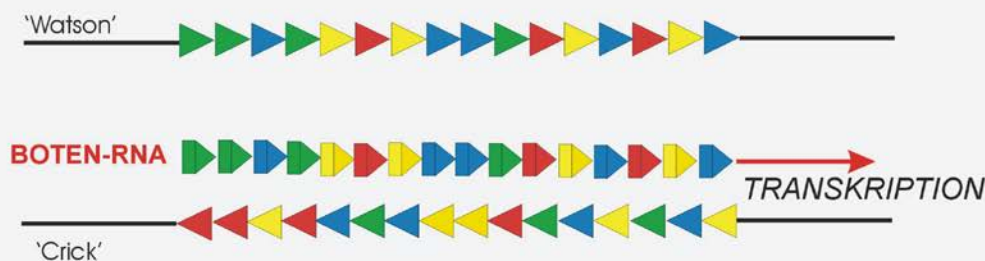
## Vom Gen zum Protein



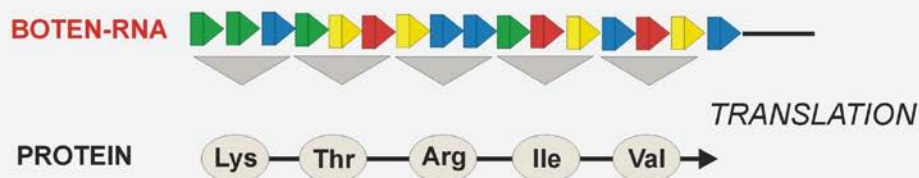
Die DNA ist ein Doppelstrang gebildet aus komplementären Basen ("Basenpaare")

Die Abfolge der Basen nennen wir die DNA-Sequenz

Ein Gen entspricht einem definierten Abschnitt der DNA-Sequenz



Zum "Gebrauch" wird die DNA-Sequenz eines Gens in eine RNA-Sequenz umkopiert, es entsteht Boten-RNA. Dazu wird ein ganzes System bestehend aus vielen Protein-Faktoren benutzt.



Die Sequenz der Boten-RNA wird mit Hilfe einer speziellen Maschinerie (Ribosomen) in eine spezifische Abfolge von Aminosäuren übersetzt, welche gleichzeitig zur Protein-Kette verknüpft werden.

Die Ablesung erfolgt in Dreier-Einheiten gemäß den Regeln des **Genetischen Codes**.

Abbildung 6: Das Zentrale Dogma aufgeschlüsselt.

Ein detaillierteres Bild können wir in Abbildung 6 zeichnen: Die in der Basenabfolge der DNA niedergelegte Information wird zunächst in „Einheiten“, die einem Gen entsprechen, mehrfach kopiert und dabei in eine verwandte Struktur (RNA, Ribonucleinsäure) übertragen; einerseits darf das genetische Material nicht angetastet werden, andererseits müssen diese Kopien in regulierter Weise erstellt und nach Gebrauch auch wieder eliminiert werden können. An diese **Transkription** schließt



sich die **Translation** an: Die Proteinsynthese-Maschinerie der Zelle benutzt die Boten- oder ‚messenger‘-RNA (kurz mRNA) als 'Vorschrift', anhand derer sie das von einem Gen 'codierte' Protein herstellt: Jeweils drei aufeinander folgende Basen (ein Basentriplett) bestimmen, welche der 20 Aminosäuren ausgewählt und an die wachsende Peptidkette geknüpft wird. Der Genetische Code, gewissermaßen die Tabelle, die jedem möglichen Basentriplett eine bestimmte Aminosäure zuordnet, erwies sich als nahezu universell für alle Organismen (Abbildung 7). Die Notwendigkeit eines durchgehenden Leserasters machte unmittelbar verständlich, dass die meisten Mutationen auf Brüche in diesem Raster zurückgehen, auch dann, wenn nur eine einzige Base an entscheidender Stelle verändert wird oder herausfällt: die Konsequenz ist in der Regel ein nicht mehr voll funktionstüchtiges oder gar ein funktionsuntüchtiges Protein. Darüber hinaus lieferte die Universalität des genetischen Codes die Basis für alle späteren Unternehmen, auch ein in einen Organismus eingeführtes ‚fremdes‘ Gen in das von ihm codierte Protein umzusetzen.

		Zweiter Buchstabe			
		U	C	A	G
Erster Buchstabe	U	UUU <b>Phe</b> UUC UUA <b>Leu</b> UUG	UCU UCC UCA <b>Ser</b> UCG	UAU <b>Tyr</b> UAC UAA <b>Stop</b> UAG	UGU <b>Cys</b> UGC UGA <b>Stop</b> UGG <b>Trp</b>
		CUU CUC <b>Leu</b> CUA CUG	CCU CCC <b>Pro</b> CCA CCG	CAU <b>His</b> CAC CAA <b>Gln</b> CAG	CGU CGC <b>Arg</b> CGA CGG
		AUU AUC <b>Ile</b> AUA AUG <b>Met</b>	ACU ACC <b>Thr</b> ACA ACG	AAU <b>Asn</b> AAC AAA <b>Lys</b> AAG	AGU <b>Ser</b> AGC AGA <b>Arg</b> AGG
		GUU GUC GUA <b>Val</b> GUG	GCU GCC GCA <b>Ala</b> GCG	GAU <b>Asp</b> GAC GAA <b>Glu</b> GAG	GGU GGC GGA <b>Gly</b> GGG

Abbildung 7: Die Standard-Version des Genetischen Codes. Alle 64 möglichen Wörter (Codons jeweils bestehend aus drei Basen) sind mit den von ihnen codierten Aminosäuren in diesem Schema angeordnet. Der erste ‚Buchstabe‘ ist die erste Base des Triplets, der zweite Buchstabe die ‚innere‘

Base des Triplets. Ihnen folgt der dritte Buchstabe, womit sich vier Codons in einem Kasten ergeben. Die 20 verschiedenen Aminosäuren sind farbig gekennzeichnet: Die ‚gelben‘ Symbole bezeichnen unpolare Aminosäuren (Phe=Phenylalanin; Leu=Leucin; Ile=Isoleucin; Val=Valin; Pro=Prolin; Ala=Alanin; Trp=Tryptophan; Gly=Glycin). Polare Aminosäuren sind hellblau (Ser=Serin; Thr=Threonin; Cys=Cystein). Basische Aminosäuren sind violett gekennzeichnet (Lys=Lysin; Arg=Arginin), orange erscheinen die sauren Aminosäuren (Asp=Asparaginsäure; Glu=Glutaminsäure). Methionin (Met), repräsentiert durch das Codon AUG, wird als Start für jede Peptidkette benutzt. Drei Codons (UAA, UAG und UGA codieren keine Aminosäure sondern bedeuten, dass die Synthese der Peptidkette hier stoppen soll. Einzelne Aminosäuren (Met, Trp) haben nur ein Codewort; andere Aminosäuren benutzen zwei, drei, vier, oder sogar sechs Codewörter. Dadurch wird u.a. die Häufigkeit ihres Einbaues in die Peptidkette bestimmt.

In den Jahren, die der Entdeckung der Doppelhelix folgten, wurden Details zu den Prozessen ausgearbeitet, die für die Weitergabe und die Umsetzung der genetischen Information verantwortlich sind.







Genomgrößen (in Basenpaaren) und Genzahl		
	Viren	<div>5 -10 000</div> <div>5 -10</div>
	Bakterien	<div>1-5 000 000</div> <div>800 - 4 000</div>
	Bierhefe	<div>13 000 000</div> <div>6 000</div>
	Drosophila	<div>100 000 000</div> <div>20 000</div>
	Maus	<div>3 000 000 000</div> <div>~20 000</div>
	Mensch	<div>3 000 000 000</div> <div>~20 000</div>

Abbildung 8: Genomgröße und Zahl der Gene

Während zunächst Methoden zur Verfügung standen, Proteine in reiner Form zu isolieren, die Abfolge ihrer Aminosäure-Bausteine und ihre dreidimensionale Struktur zu bestimmen, und sogar kleinere RNA-Moleküle zu isolieren und ihre Struktur aufzuklären, gab es noch keine Möglichkeit, einzelne Gene in die Hand zu bekommen. Man muss bedenken, dass jede Zelle nur *eine* Kopie der genomischen DNA, also äußerst wenig DNA-Material, enthält. Außerdem, wie sollte man ein einzelnes Gen aus einem ‚Riesenmolekül‘, wie einem ganzen Chromosom, herausoperieren? Eine Rechnung mit einfachen Annahmen veranschaulicht das Problem. Ein Gen, welches ein Protein aus 200 Aminosäuren

codiert, braucht mindestens 600 Basenpaare an Information. Hätte ein Organismus nur 1000 Gene, so ergäbe das eine Mindestlänge der genomischen DNA von 600.000 Basenpaaren (unter der Annahme, dass jeweils nur einer der beiden DNA-Stränge codiert). Wie sich herausstellte, müssen aber zusätzlich noch Zeichen für Anfang und Ende der Botschaft und für weitere Signale untergebracht werden. Niemand konnte anfangs also die tatsächliche Länge der DNA für ein ihn interessierendes Gen nicht im Entferntesten abschätzen (Abbildung 8).

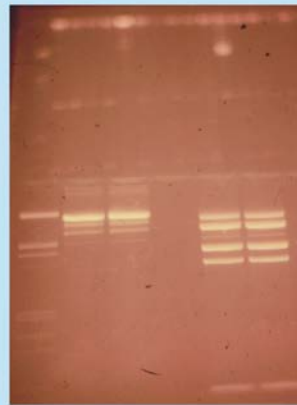
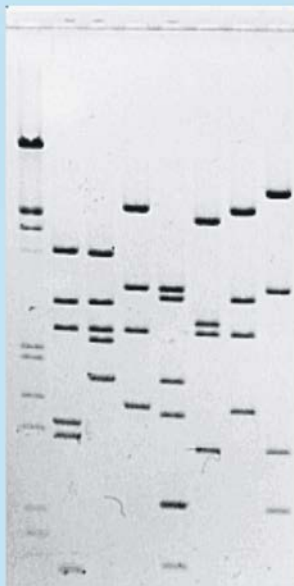
#### 1.4 Rekombinante DNA - Gene in Substanz

Die Jahre zwischen 1972 und 1977 kommen einer **molekularbiologischen Revolution** gleich [1]. Es waren drei wesentliche Entwicklungen, die entscheidend dazu beitrugen, dass man Gene wirklich in Substanz fassen konnte: Die Entdeckung der **Restriktionsenzyme** (Abbildung 9); die Darstellung **rekombinanter DNA** (Abbildung 10); und die Etablierung von Methoden zur **Sequenzierung von DNA** (Abbildung 11).



Werner Arber 2008

## Restriktionsenzyme



Auftrennung der DNA-Fragmente in der Gelelektrophorese

Abbildung 9: Wirkungsweise eines von über 200 heute bekannten und in der Molekularbiologie angewandten Restriktions-Enzyme. Jedes dieser Enzyme besitzt eine Erkennungssequenz (Hier GGATCC) innerhalb der DNA-Sequenz. Durch Schnitte an diesen Stellen wird die DNA in Fragmente zerlegt, welche durch sog. Gelelektrophorese nach ihrer Größe aufgetrennt werden (links sichtbar gemacht durch radioaktive Markierung); jede Bahn stammt aus einer solchen Spaltung. Man kann die Fragmente auch durch Färbung oder unter Fluoreszenzlicht sichtbar machen (rechts).

Aus der Bakteriengenetik wusste man, dass genetische Information natürlicherweise von einem ‚Spender‘ auf einen ‚Empfänger‘ übertragen und umgesetzt werden kann. Das eklatanteste Beispiel dafür liefern pathogene Bakterien, die Resistenzen gegen bestimmte Antibiotika erwerben können, ein Phänomen, das seit 1953 als Hospitalismus bekannt ist. Die Resistenzen werden durch Gene vermittelt, die auf kleinen ringförmigen DNA-Molekülen (Plasmiden) außerhalb des eigentlichen Genoms in vielen Kopien bewahrt werden. Außerdem können Plasmide leicht von einer Zelle auf eine andere übertragen werden. Hat ein Bakterium es einmal geschafft, sich ein bestimmtes Resistenzgen einzuverleiben, so hat es zugleich die Widerstandskraft gegen das entsprechende Antibiotikum erworben. In der heutigen Zeit hat der oft unüberlegte Einsatz von Antibiotika dazu geführt, dass Bakterien viele Resistenzen entwickeln, die sich ansammeln und auf andere Keime übertragen werden können.

Die Existenz des geschilderten Mechanismus zur Übertragung genetischer Information von einem Bakterium auf ein anderes führte Paul Berg 1971 als ersten zu einer entscheidenden Frage: Kann man auch ein DNA-Molekül tierischer Provenienz in eine Bakterienzelle als ‚Wirt‘ einschleusen, und

wird dessen Information in der neuen Umgebung auch umgesetzt? Er verknüpfte dazu ein Stück bakterieller DNA mit einem Stück der fremden DNA. Diese neue Kombination von zwei verschiedenen DNA-Molekülen, eben eine **rekombinante DNA**, ließ sich wie ein Plasmid in das Bakterium *Escherichia coli* (ein obligater Bewohner unserer Darmflora und von Beginn an **das** Modell der Bakteriengenetik) einführen - und das Experiment gelang [2].

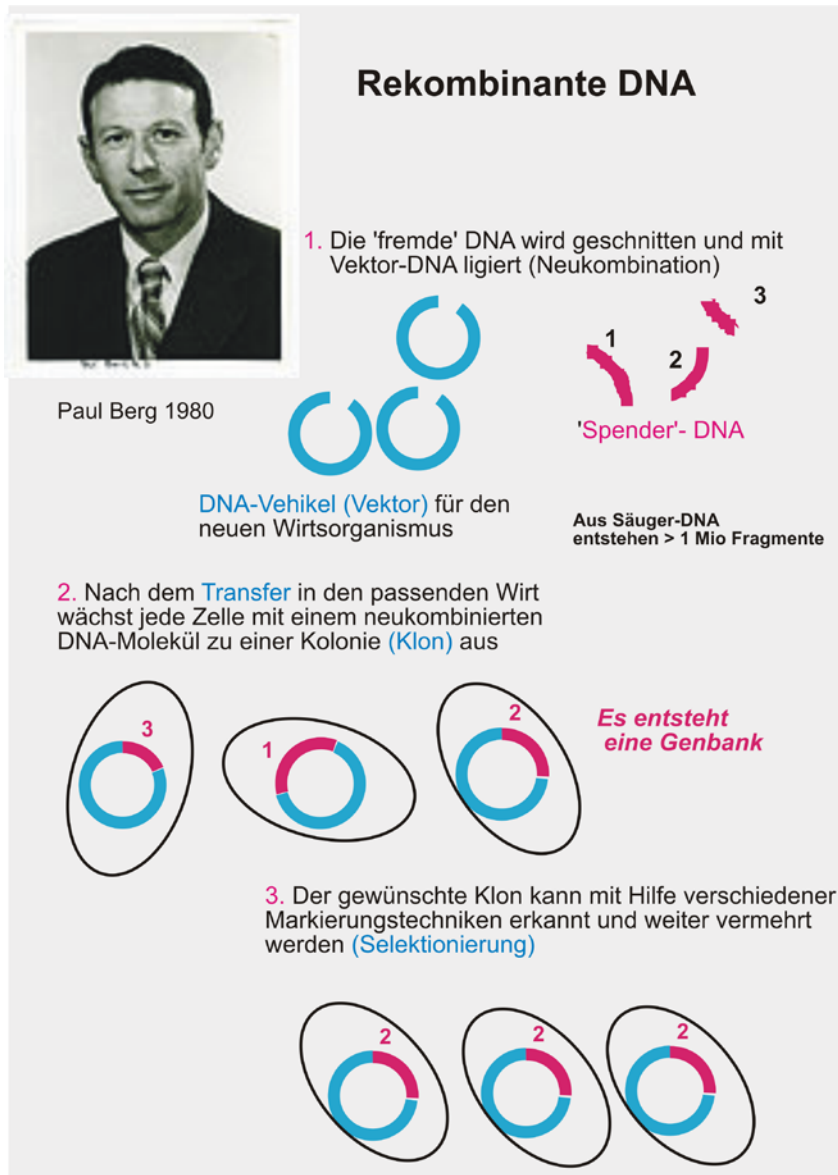


Abbildung 10: Molekulares Klonen.

Trotz der unangenehmen Eigenschaft mancher Plasmide Resistenzgene (das sind Gene, die Bakterien u.a. gegen Antibiotika unempfindlich machen) weiterzugeben, boten sie sich als fast ideale und generell verwendbare Vehikel an, um rekombinante DNA in Bakterien, insbesondere *E. coli*, einzuschleusen. Den führenden Forschern war durchaus bewusst, dass solches Vorgehen ein gewisses Wagnis war, und so verlangten sie 1974, zunächst alle Experimente mit rekombinanter DNA von Viren, Toxin- oder Resistenzgenen auszusetzen [3]. Um aber die immensen Chancen der neuen Technik nicht zu verspielen, einigte man sich im Frühjahr 1975 auf der von der *National Academy of*



*Sciences* nach Asilomar einberufenen Konferenz darauf, in Verbindung mit den staatlichen Sicherheitsbehörden genaue Regeln aufzustellen, an die sich die Forscher bei Experimenten mit rekombinanter DNA zu halten hatten [4]. Um alle Risiken auszuschließen, mussten nicht nur besondere Sicherheitsvorkehrungen eingehalten, sondern auch ‚entschärfte‘ Bakterienstämme und Plasmide verwendet werden, die anhand detaillierter Experimente entwickelt worden waren und nur im Labor, aber nicht in der freien Natur, zu überleben vermochten. Aus diesen Richtlinien sind dann später ländereigene Gesetze zum Umgang mit gentechnisch veränderten Organismen (in Deutschland die "Gentechnik-Gesetze") hervorgegangen. Unter den festgelegten und von allen Beteiligten eingehaltenen Sicherheitsvorkehrungen sind - auch in der bald anlaufenden industriellen Produktion rekombinanter Genprodukte - in den fast vierzig folgenden Jahren keine Unfälle aufgetreten.

Von unschätzbarem Vorteil erwiesen sich bei der Darstellung rekombinanter DNA die so genannten Restriktionsenzyme, die Werner Arber kurz zuvor entdeckt hatte und von denen heute weit über 200 mit verschiedener Spezifität kommerziell erhältlich sind. Diese, aus Bakterien isolierten, Enzyme vermögen ein DNA-Molekül an einer für jedes Enzym charakteristischen Sequenz (die je nach Enzym vier bis zu zehn Basenpaare umfasst) zu schneiden und so jede DNA, also selbst genomische DNA, in definierte Fragmente zu zerlegen, die dann alle über die ‚gleichen‘ Enden verfügen. Unter bestimmten Reaktionsbedingungen kann jedes dieser Fragmente in einem 'Eintopf-Verfahren' mit einem geeigneten kleinen DNA-Molekül (zum Beispiel einem ‚entschärften‘ Plasmid) als Vehikel verknüpft und das Gemisch einer Kultur von Wirtszellen, welche viele Millionen Zellen enthält, angeboten werden. Die Genetiker hatten gelernt, dass eine einzelne Wirtszelle immer nur ein einziges solches Molekül aufnehmen kann. Lässt man die Kultur wachsen, so vermehrt jede Zelle also auch nur das eine gerade von ihm aufgenommene Molekül: Sie repräsentiert einen ‚Klon‘, der ein ‚kloniertes‘ DNA-Fragment enthält. Insgesamt liegt nun eine Klon-Bibliothek vor, aus der man dann einzelne Klone mit einem bestimmten DNA-Fragment, z.B. einem gewünschten Gen, auswählen kann. Dazu wurde im Laufe der Zeit eine Vielzahl geeigneter Erkennungs- und Selektionsverfahren entwickelt. Der bestechendste Aspekt der Klonierung war jedoch die Tatsache, dass sich nun ein bakterieller Klon und mit ihm das eingeschleuste DNA-Fragment durch Aufzucht Millionen Mal vervielfachen ließ.

Die Verfügbarkeit von DNA in Substanz verlangte jetzt auch Techniken, mit denen man rasch und sicher die Basenabfolge eines DNA-Moleküls ermitteln konnte. Eine von *Maxam* und *Gilbert* erdachte chemische Methode erwies sich hierfür zwar als nützlich und zuverlässig, aber letztlich als zu aufwändig. *Frederick Sanger* ging das Problem der DNA-Sequenzierung von einer anderen Seite an und entwickelte 1977 eine elegante Methode [5], bei der das DNA-Fragment, dessen Basenabfolge es zu ermitteln gilt, viele hunderttausend Male im Reagenzglas durch zugesetzte Enzyme kopiert wird, aber so, dass vom gleichen Ende aus immer nur kürzere Moleküle entstehen, deren letzte Base außerdem noch durch eine Markierung – früher durch radioaktive Markierung, neuerdings durch eine fluoreszierende Base, gekennzeichnet wird. Das Gemisch wird dann durch eine so genannte Gelelektrophorese getrennt, wobei die rein statistisch entstandenen Moleküle der Länge nach sortiert

werden - jedes Molekül ist genau um eine Base länger oder kürzer als das 'benachbarte', und man erhält eine Art Leiter. (In Abbildung 11 ist das Radioautogramm einer Gelelektrophorese gezeigt; die Basensequenz lässt sich ablesen, wenn man jeweils zum nächsten „Strich“ in einer der vier Leitern weitergeht, siehe rechte Seite auf dem Radioautogramm).

Hat man bei der Markierung für jede der vier Basen eine bestimmte Fluoreszenz-Farbe gewählt, so lässt sich aus der Farbe der Leitersprossen die Sequenz unmittelbar ablesen. Die Sanger-Technik hat sich als erfolgreichste und schnellste Methode der DNA-Sequenzierung durchgesetzt und bildet heute die Grundlage für viele der automatisierten Verfahren der DNA-Analyse (Abbildung 12).



Abbildung 11: Radioautogramm einer DNA-Sequenz nach der Sanger'schen Didesoxy-Kettenabbruch-Methode.

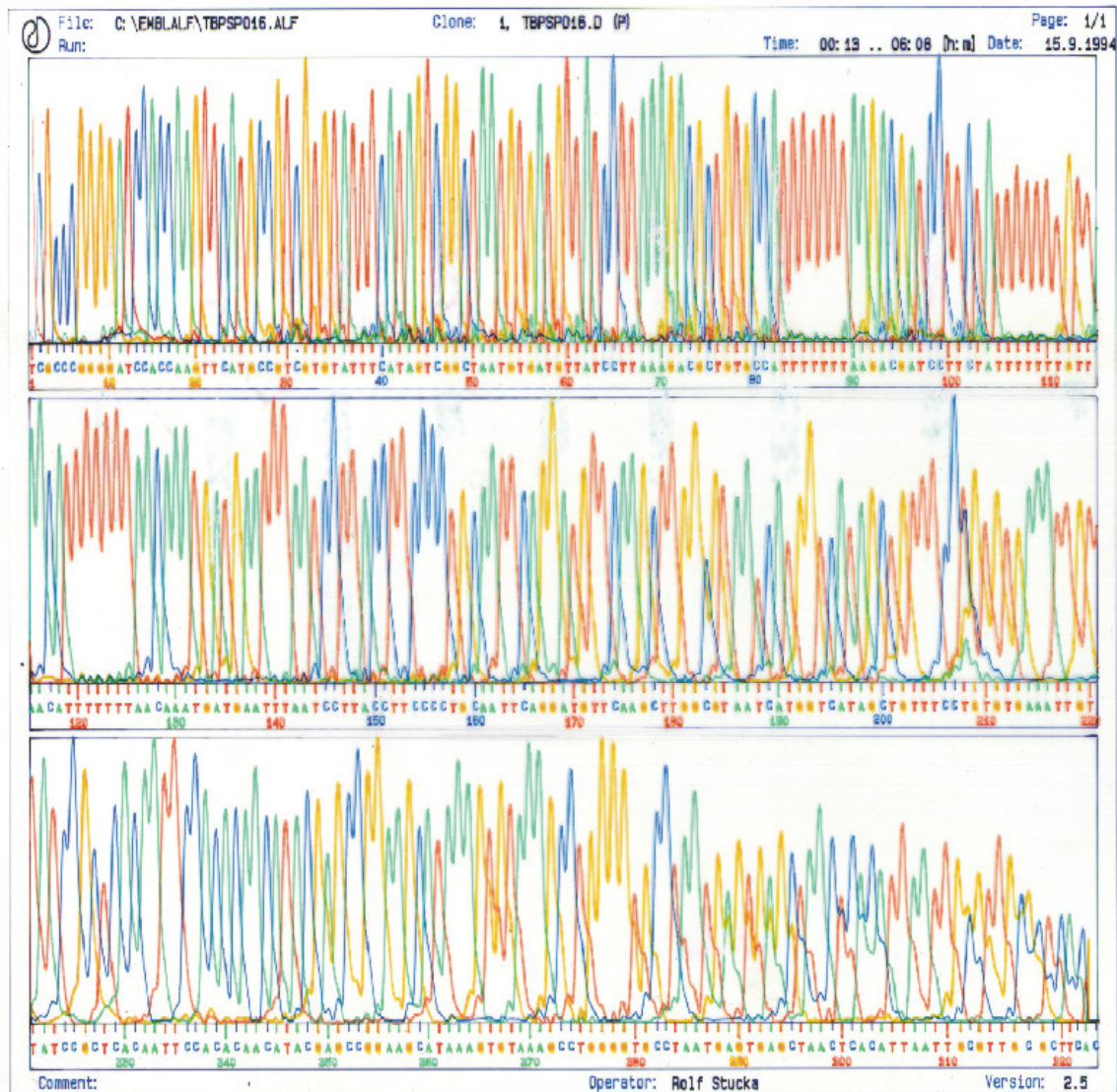


Abbildung 12: Auslesung einer DNA-Sequenz, erstellt durch Gelelektrophorese mit vier verschiedenfarbig markierten ‚Kettenabbruch-Reagenzien‘ (EMBL Automat nach W. Ansorge, Heidelberg).

Die Weiterentwicklung von **Sequenzieretechniken** ließ nicht lange auf sich warten; es kam besonders darauf an, verlässliche Daten schnell und von sehr großen Genomen zu erhalten. Zunächst zur Entschlüsselung mikrobieller Genome, dann selbst im Humangenomprojekt (siehe Abschnitt **2.3**) wurde die sogenannte *shotgun*-Technik eingesetzt. Diese beinhaltete die Erzeugung von Millionen etwa 400 Basenpaaren langen und zufällig entstandenen DNA-Fragmenten aus dem Gesamtgenom, die Sequenzanalyse aller dieser Fragmente, sowie die Zusammensetzung der Gesamtsequenz aus überlappenden Teilstücken mit Hilfe von Computermethoden.

Schon im Verlaufe des Humangenomprojektes kamen neuere und schnellere Sequenzierautomaten auf den Markt und wurden in manchen Labors zu Hunderten nebeneinander eingesetzt. Parallel dazu erfolgte auch die Entwicklung neuer Techniken. In der englischen Wikipedia findet man unter dem

Stichwort ‚DNA sequencing‘ die Beschreibung von mindestens zehn verschiedenen Verfahren des „Next-generation sequencing“ (und deren Vergleich [6]).

Mindestens acht unterschiedliche Ansätze (‚third-generation sequencing‘) werden dort unter „methods in development“ beschrieben, die aber zum Teil noch ausreifen müssen. Zum Beispiel werden in einem Verfahren DNA-Fragmente an eine biologische Nanopore gekoppelt und durch diese vorwärts bewegt: Sobald die Nukleotide durch die Pore treten, ändert sich der elektrische Fluss; jede der vier Basen ergibt ein unterscheidbares Signal, welches ausgelesen wird [7]. Ein ähnliches Verfahren zur Sequenzierung von einzelsträngiger DNA benutzt die Messung elektrischer Tunnelströme, wenn sich das DNA-Molekül durch einen Kanal bewegt; jede Base bewirkt eine charakteristische Veränderung des Tunnelstromes [8,9].

Zu diesen Sequenzier-Techniken hat sich seit 1980 ein sehr schnelles und effizientes Verfahren zur ‚wunderbaren DNA-Vermehrung‘ hinzu gesellt, die so genannte **Polymerase-Kettenreaktion (PCR)**, welche von Kary Mullis in den USA entwickelt wurde [Abbildung 13, siehe auch ref. 10]. Heute stehen über Hundert Protokolle für die PCR zur Verfügung, und immer ausgeklügeltere Automaten werden auf dem Markt angeboten.

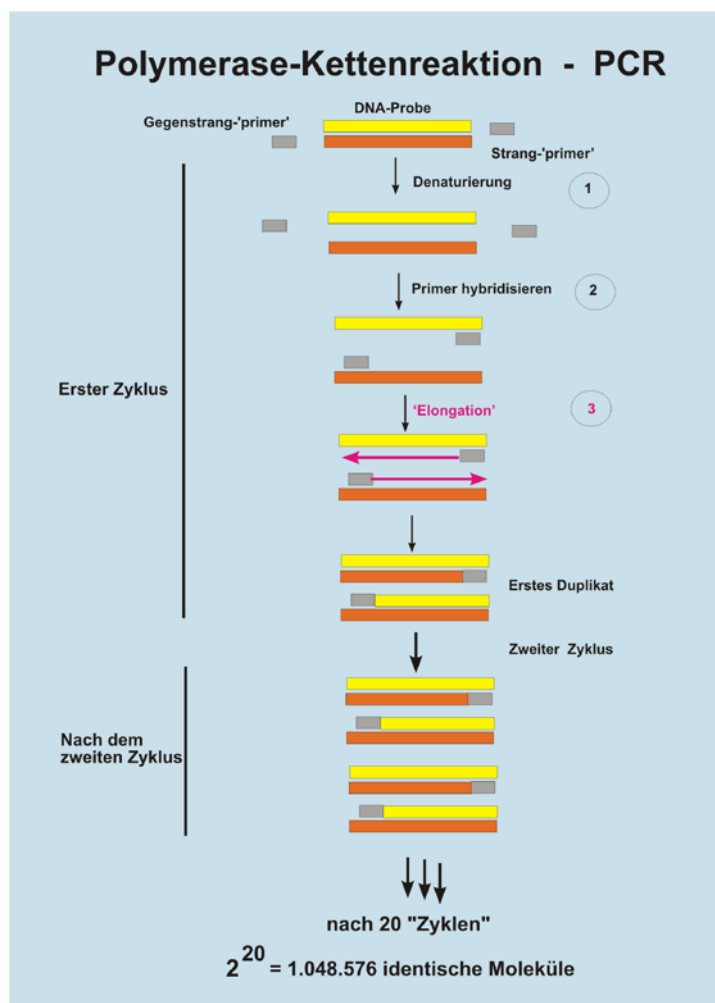


Abbildung 13: Prinzip der Polymerase-Kettenreaktion. Mit Hilfe dieser 1977 entwickelten Technik können bis zu 40000 Basenpaare von DNA-Fragmenten vervielfältigt werden. Man benötigt für die Reaktion zwei ‚primer‘ (kurze komplementierende DNA-Stücke), alle Desoxytrinukleotide und eine hitzestabile DNA-Polymerase, welche die Stränge durch Synthese vervollständigt. Die Ansätze werden in Mikrogefäßen gefahren und alle Reaktionsschritte mit Hilfe eines sog. Thermozyklers gesteuert. Im ersten Schritt erfolgt durch Erhitzen der Probe auf 96° die Trennung der DNA-Stränge (und der ‚primer‘); im zweiten Schritt wird auf 55-65° heruntergekühlt, so dass sich die ‚primer‘ anlagern können. Im dritten Schritt (bei 68-72°) lässt man Zeit, damit die DNA-Polymerase die



*beiden Stränge kopieren kann; sie bleiben getrennt. In den weiteren Zyklen werden diese drei Schritte prinzipiell wiederholt. Nach etwa 30 Zyklen hat sich infolge exponentieller Vermehrung der eingesetzten DNA genügend Material für die weitere Verwendung angesammelt.*

Mit der **PCR-Technik** kann man von einer DNA-Vorlage Millionen von gleichartigen Kopien gewissermaßen im Reagenzglas erzeugen, um ausreichend Material für die weitere Analyse zu gewinnen. Die PCR ist heute vor allem unentbehrlich für die Bereitstellung von sequenzierbarem DNA-Material. Sie wurde und wird auch in der so genannten DNA-Diagnostik eingesetzt, bei der man geeignete Sonden (DNA-Fragmente) verwendet, um mutierte Gene oder in den Körper eingedrungene pathogene Keime aufzuspüren. Nicht mehr wegzudenken ist diese Methode in der pränatalen Diagnostik, und sie hat auch seit ihrer ersten Anwendung in den 80er Jahren jetzt ihren festen Platz in der forensischen Medizin, um Personen anhand individueller DNA-Sequenzen zu identifizieren. Als Ausgangsmaterial genügt dabei ein einziges DNA-Molekül, das sich beispielsweise aus einer Blutprobe, einem Stückchen Knochen, aus einer Haarwurzel oder Spermaspuren isolieren lässt.

## **1.5 Reverse („Umgekehrte“) Genetik**

Während anfangs in den meisten Fällen, in denen die biologische Information genutzt werden sollte, Gene aufgrund ihrer Funktion gesucht und kloniert wurden, beschritten die Forscher in den 80er Jahren den umgekehrten Weg: Sie begannen, DNA aus verschiedenen Organismen systematisch zu sequenzieren, um mit Hilfe computer-gesteuerter Programme Gene vorherzusagen und diese dann auf ihre Funktion hin zu analysieren. Da die Sequenziertechnik damals die Bestimmung nur jeweils relativ kurzer Sequenzen (in der Größenordnung von einigen hundert Basenpaaren) erlaubte, wurden gleichzeitig Strategien etabliert, um die Sequenz längerer DNA-Moleküle aus 'überlappenden' DNA-Fragmenten zusammenzusetzen. Als geeignete Objekte wählte man zunächst DNA-Moleküle von überschaubarer Größe aus, wie zum Beispiel die Genome von Mitochondrien und Viren in der Größenordnung von einigen tausend Basenpaaren (KB = ‚Kilobasen‘; vgl. dagegen MB = ‚Megabasen‘ bedeutet Millionen Basenpaare).

Heute, wo die Sequenzermittlung praktisch keine technischen oder zeitlichen Grenzen mehr kennt, sondern vielleicht nur noch eine Kostenfrage darstellt, gibt es mittlerweile unzählige auf Sequenzierung basierende Projekte, welche von ganzen Konsortien bearbeitet werden oder die im Routinefall auch eigens dafür etablierte Firmen übernehmen [11].

Im Folgenden werden wir uns mit dieser Entwicklung näher befassen.

## 2 Genomprojekte

### 2.1 Gene und Genome

Die rasche Weiterentwicklung der Sequenzier-Techniken und die Einführung von automatisierten Verfahren der DNA-Analyse ließ es Ende der 80er Jahre möglich erscheinen, auch Genome von ganzen Organismen und sogar das menschliche Genom zu sequenzieren, eines der ehrgeizigsten Genomprojekte, über das noch eingehender zu sprechen sein wird. Die im Jahre 1989 für das Humangenomprojekt veranschlagte Laufzeit von 15 Jahren erschien den Forschern nützlich, sich zunächst mit Genomen weniger komplexer Organismen, quasi als ‚Modellsystemen‘, zu befassen, um an ihnen Techniken zu vervollkommen und vor allem Einsichten in Genstrukturen und die Anlage von Genomen zu gewinnen. In der Tat konnten hier wertvolle Erkenntnisse gesammelt werden, zumal man auch bewusst Organismen auswählte, die entweder neue Erkenntnisse zur Evolution bestimmter Spezies oder Einsichten in Pathogenitätsmechanismen von Krankheitserregern versprachen. Das erste Genom eines pathogenen Bakteriums, *Haemophilus influenzae*, wurde 1995 entschlüsselt. Unzählige Genome gefährlicher Keime folgten (siehe Abschnitt 2.3).

### 2.2 Die Bäckerhefe als eukaryotisches Modellsystem

#### 2.2.1 Das Hefegenom wird sequenziert

Die Hefeforscher fanden ihr ‚Haustier‘ für ein international koordinierbares Programm, das im Jahre 1989 initiiert wurde, besonders attraktiv: Die **Bäcker- oder Bier-Hefe** (*Saccharomyces cerevisiae*) ist ein kernhaltiger, einzelliger Mikroorganismus, rasch wachsend und leicht zu handhaben. Sie war zu dieser Zeit genetisch und biochemisch bereits eingehend charakterisiert und wurde zudem seit tausenden von Jahren als unerlässlicher Helfer in der Back- und Brautechnik kultiviert. Hefe darf aber auch als Prototyp für Pilze gelten, die sich von jeher als interessante experimentelle Organismen anboten, weil sie sowohl Produzenten von wertvollen Antibiotika oder anderen Biomolekülen darstellen, aber auch als Pathogene von Pflanzen und Tieren auftreten.

Das Hefegenomprojekt startete mit finanzieller Unterstützung durch die Biotechnologie-Programme der Europäischen Union. In Zusammenarbeit von 35 Gruppen aus europäischen Laboratorien wurde 1992 die vollständige Sequenz des Hefechromosoms III, und damit die erste Sequenz eines eukaryotischen Chromosoms überhaupt, vorgelegt. Weitere Chromosomen folgten zwei Jahre später. Gruppen aus den USA, Kanada und Japan mit insgesamt 600 Wissenschaftlern beteiligten sich schließlich an diesem Projekt, und Anfang 1996 war die gesamte Sequenz des Hefegenoms (13 Megabasen = 13 Mio. Basenpaare) entschlüsselt [12]. Die Analyse der Daten ergab, dass die Hefezelle etwa 6.000 Gene beherbergt, welche sie zu allen Lebensfunktionen befähigen [13]. Neuere Daten findet der Leser in einer Monografie über Hefe [14].

### 2.2.2 Analyse von Genfunktionen

Vor Beginn des Hefegenomprojektes hatten Hefegenetiker bereits die Funktionen von etwa 1200 Hefegenen festlegen können. Man war dabei in ‚klassischer‘ Weise so vorgegangen, dass man Hefemutanten erzeugte, die eine bestimmte Funktion verloren hatten. Die Lage des zugehörigen Gens konnte durch Kreuzungsanalyse bestimmt, anschließend das Gen selbst isoliert und seine Sequenz ermittelt werden. Durch Rückübersetzung der DNA-Sequenz mit Hilfe des genetischen Codes erhielt man die Aminosäuresequenz des betreffenden Proteins.

Ähnlich verfuhr man mit zahlreichen Genen aus anderen genetisch gut charakterisierten Organismen, wie zum Beispiel dem Bakterium *Escherichia coli* oder der Taufliege *Drosophila*, man hatte sogar die Funktionen einer beträchtlichen Zahl menschlicher Gene ermitteln können. Diese Daten wurden ab etwa 1980 in allgemein zugänglichen Datenbanken gesammelt und bildeten die Grundlage für computergestützte Vergleiche aller jeweils verfügbaren DNA- und Proteinsequenzen.

Die vergleichende Analyse - heute sagt man *in silico* Analyse - liefert Hinweise sowohl auf die Verwandtschaft von Proteinen in einem Organismus als auch auf die Verwandtschaft von Proteinen aus verschiedenen Organismen und gestattet ihre evolutionsgeschichtliche Einordnung: Proteine aus verschiedenen Organismen weisen umso signifikantere Ähnlichkeiten in ihrer Sequenz aus, je näher diese in der Evolution einander stehen. Auf diese Weise lassen sich ‚homologe‘ Proteine erkennen, und man kann aus Sequenzvergleichen dann auf bislang in einem Organismus nicht bekannte Genfunktionen zurückschließen, wenn diese in einem anderen Organismus bereits identifiziert wurden.

So konnte man nun auch systematisch nach Verwandtschaften unter allen Hefegenen und den von ihnen codierten Proteinen suchen. Da man davon ausgehen kann, dass Gene, welche besonders eng verwandt sind, sich aus einem gemeinsamen Vorläufer entwickelt haben und deshalb ähnliche Funktionen ausüben, konnten die Hefegene in einzelne Genfamilien geordnet werden; in manchen Fällen umfassen solche Genfamilien bis zu 50 oder mehr Mitglieder.

Trotz der stetig (fast exponentiell) anwachsenden Zahl von Datenbankeinträgen lässt sich aus systematischen Vergleichen bis heute mehr als 10% aller Hefegene auch nicht annäherungsweise eine Funktion zuschreiben. Dieses Problem hat sich überraschenderweise auch bei der vollständigen Sequenzierung der Genome von allen weiteren, auch den bakteriellen, Organismen ergeben: im Mittel bleiben etwa 40% der Gene unbekannter Funktion. Zur Erklärung muss man annehmen, dass diese Gene vom jeweiligen Organismus individuell entwickelt wurden, um sich besonderen Lebens- oder Umweltbedingungen anzupassen. Hier werden also letztlich nur gezielte biochemische Analysen Aufklärung bringen. Eine weitere Überraschung findet sich in der Tatsache, dass in allen bisher sequenzierten Genomen ein großer Teil der Gene (ca. 30-40%) ‚doppelt‘ angelegt ist. Sehr verkürzt formuliert, gibt es mehrere mögliche Erklärungen dafür: (1) Die doppelte Anlage bietet Schutz vor Verlust, (2) das ‚zweite‘ Gen wird nur unter besonderen Umwelt-Bedingungen aktiviert, (3) eine Kopie dient als ‚Spielmaterial‘ in der Evolution.

### 2.2.3 Im Hefegenom finden sich menschliche ‚Krankheitsgene‘

Bei systematischen Abfragen der in den Datenbanken vorhandenen Protein-Einträge stellte sich auch ein beträchtlich hoher Verwandtschaftsgrad von vielen Hefegenen mit Humangenomen heraus, insbesondere solchen, die für Basisfunktionen des Zellstoffwechsels zuständig sind: Gene, welche die Zellteilung und die dafür notwendige Duplikation der DNA überwachen; Gene, die eine koordinierte Umsetzung der Geninformation ("zur rechten Zeit am rechten Ort") steuern; Gene, die den Informationsaustausch zwischen Zellen kontrollieren oder Enzymaktivitäten vermitteln, um aufgenommene Nährstoffe in zelleigene Bestandteile umzusetzen und die dafür erforderliche Energie bereitzustellen. Mit anderen Worten, Prinzipien, die sich sozusagen in einem primitiven Eukaryoten bewährt hatten, wurden durch die gesamte Evolution beibehalten. Am eindringlichsten wird das dadurch veranschaulicht, dass man die Funktion von solchen Hefegenen durch das Einschleusen der homologen Humangene kompensieren kann. Darüber hinaus konnten in der Hefe sofort über 100 Pendanten zu menschlichen ‚Krankheitsgenen‘ identifiziert werden, also Genen, die in mutierter Form zu erblichen Anomalien führen. Mit Hilfe der Sequenzen von Hefegenen konnten sogar viele bis dahin unbekannte Krankheitsgene identifiziert werden (heute über Tausend). Darüber hinaus eröffnete sich die Möglichkeit, die Funktionen unbekannter Humangene zu analysieren, wenn man sie in die Hefe einbringt, weil das Hefesystem eine Reihe von einfachen Techniken dafür bietet (siehe in ref.14).

### 2.2.4 Hefen als Modellsysteme für die Evolution

Im Jahre 2000 berichteten französische Hefe-Forscher über erste Ergebnisse [15] der Sequenzierung von 13 verschiedenen Hefen (Génolevures Programme), die sich sowohl im Hinblick auf ihre Eigenschaften als auch auf ihre evolutionäre Verwandtschaft als aussichtsreiche Modelle anboten. Bezeichnenderweise gehörten die ausgewählten Spezies zu den sogenannten Hemiascomyceten, die zum Teil sehr unterschiedliche Stoffwechselwege beschreiten (Tabelle 2.1).

Tabelle 2.1: Charakteristika einiger Hefe-Spezies

Spezies	Vorkommen/Charakteristika	Anwendungen
<i>Saccharomyces bayanus</i>		Weinproduktion
<i>Saccharomyces exiguus</i>	Im Boden, in Abwässern	Fermentieren von ‚softdrinks‘
<i>Saccharomyces zervazii</i>		Fermentieren von Trehalose und Glycerin
<i>Candida glabrata</i>	Menschliches Pathogen	
<i>Zygosaccharomyces rouxii</i>	Salztolerant, osmotolerant	Produktion von Sojasauce
<i>Saccharomyces kluyveri</i>	Wächst auf Pyrimidinen	
<i>Kluyveromyces lactis</i>	Wächst fermentativ oder aerob	Milchindustrie, für biotechnische Produkte
<i>Kluyveromyces thermotolerans</i>	Wächst anaerob	
<i>Kluyveromyces marxianus</i>	Inulin umsetzend	Manchmal infektiös
<i>Pichia angusta</i>	Methylotrope Hefe	Rekombinante Pharmaka
<i>Debaromyces hansenii</i>	Meeresbewohner, Kälte- und osmotolerant	Assimiliert Kohlenwasserstoffe
<i>Pichia sorbitophila</i>	Wächst unter hohem osmotischem Druck (<4 M NaCl)	
<i>Pichia pastoris</i>	Methylotrope Hefe	Methanol als C-Quelle
<i>Pichia stipidis</i>		Verstoffwechselt Xylose, Hemizellulose
<i>Candida utilis</i>	Wächst auf billigen Substraten	Produktion von Tiernahrung
<i>Candida tropicalis</i>	Menschliches, thermotolerantes Pathogen	
<i>Arxula adeninovorans</i>	Thermotolerant	Verstoffwechselt u.a. Nitrate, Phenol, Hydroxybenzoesäuren
<i>Yarrowia lipolytica</i>	Nicht-pathogen. Bemerkenswerte Sekretionswege und Lipolyse	Wächst auf Kohlenwasserstoffen, Lipiden



Später wurden andere Ascomyceten einbezogen, und bis zum Abschluss der Génolevures Programme zum Ende des Jahres 2013 wurden nun die vollständig annotierten Genome von 40 Ascomyceten vergleichbar. Wichtige Erkenntnisse hieraus betreffen erstens die Evolution dieser ‚Untergruppe‘ von Pilzen, die sich über einen Zeitraum von mehr als 1 Mrd. Jahren erstreckt, etwa der gleiche Zeitraum, in dem sich die gesamten Chordaten entwickelt haben. Zweitens lässt sich die Vielfalt der genetischen Ausstattung der Ascomyceten durch verschiedene Mechanismen erklären: Es gab und gibt Neuentstehung oder Abstoßung von Genen innerhalb einer ‚Spezies‘ – wodurch neue Varianten hervorgebracht werden -, eine Veränderung des Genetischen Codes, einen Austausch von Genen zwischen Spezies, und sogar Erwerb von Genen aus anderen Organismen (hauptsächlich aus Bakterien) [14, 16].

### 2.3 Genome von anderen Mikroorganismen

Parallel zum Hefegenomprojekt wurden Anfang und Mitte der 90er Jahre eine ganze Reihe von Projekten zur Genomsequenzierung gestartet [17, 18]). Eine neue Strategie, das so genannte ‚shot-gun‘-Sequenzieren der **Genome von Mikroorganismen** (Abbildung 14), deren Größe nur wenige Megabasen beträgt, wurde 1995 an *Haemophilus influenzae*, dem Erreger von Mittelohrentzündung und Infektionen der Atemwege, vom *Institute for Genome Research* (TIGR) etabliert. Das Prinzip bestand darin, die gesamte genomische DNA in zufällige kleine Fragmente (von etwa 400-500 BP) zu zerstückeln und Moleküle aus dieser Sammlung ‚blindlings‘ solange zu sequenzieren, bis man computerunterstützt alle Sequenzen überlappend zusammensetzen kann. Im gleichen Jahr folgten die Genome von *Mycoplasma genitalium*, einem parasitären Bakterium von Epithelzellen der Atemwege und des Genitaltraktes, und *Mycoplasma pneumoniae*, dem Verursacher atypischer Lungenentzündungen. Wegen der Adaptation an ihre ‚Umwelt‘ benötigen Mycoplasmen keine Zellwand wie andere Bakterien und können auf die Synthese von vielen Nährstoffen verzichten. So sind diese Genome mit 600 bis 800 Genen wohl die kleinsten sich selbst vermehrender Organismen.

1996 kam das Genom des ersten ‚extremophilen‘ Organismus, *Methanococcus jannaschii*, hinzu. Als Extremophile werden Mikroorganismen bezeichnet, die sich an besonders drastische Bedingungen adaptiert haben und meist auch nur unter diesen leben können: extrem hohe Temperaturen, hohe Drucke, extrem saures Milieu, hohe Salzkonzentrationen, Sauerstoff-unabhängige Energieproduktion. Extremophile gehören teils zu den echten Bakterien, teils zu den Archebakterien, welche ein eigenes Reich bilden und evolutionsgeschichtlich zwischen den echten Bakterien (Eubakterien) und den Eukaryoten stehen, weil man Charakteristika aus beiden Reichen bei ihnen findet. Die ungewöhnlichsten Extremophilen wurden aus submarinen Quellen isoliert: Sie brauchen mindestens 100°C für optimales Wachstum, aber niemand kann bisher erklären, warum Nukleinsäuren und Proteine diese Temperaturen überhaupt überstehen können, weil es keine augenfälligen Besonderheiten in deren Strukturen gibt.

## Genome pathogener Mikroorganismen

### ***Haemophilus influenzae* - 1.83 MB**

Erstes Genom eines pathogenen Bakteriums (1995).  
Erreger von Mittelohrentzündung, Infektionen der Atemwege, Meningitis

### ***Mycoplasma genitalium* - 0.6 MB - 600 Gene**

Kleinstes, sich selbst vermehrender Organismus, ohne Zellwand, obligat intrazellulär.  
Parasit von Epithelzellen der Atemwege und des Urogenitaltraktes.

### ***Mycoplasma pneumoniae* - 0.81 MB - 800 Gene**

Zellwandloser Parasit. Parasit von Epithelzellen der Atemwege und des Urogenitaltraktes.  
Menschliches Pathogen, verursacht atypische Lungenentzündungen.

### ***Helicobacter pylori* - 1.66 MB**

Flagellierte Spirochäte. Lebt als Parasit in der Magenschleimhaut.  
Verursacht chronische Entzündungen (Gastritis) und findet sich bei etwa 50% aller Menschen.

### ***Borrelia burgdorferi* - 1.44 MB**

Durch Zeckenbisse übertragene Spirochäte.  
Erreger der Borreliose (Lyme Disease)

### ***Mycobacterium tuberculosis* - 4.40 MB**

Obligat aerobes Stäbchenbakterium. Erreger der Tuberkulose, induziert granulomatöse Entzündungsreaktionen und Kavernenbildung.

### ***Treponema pallidum* - 1.14 MB**

Spirochäte, Erreger der Syphilis. Bisher keine Virulenzfaktoren und kein Impfstoff bekannt.

### ***Chlamydia trachomatis* - 1.05 MB**

Obligates Pathogen. Nicht in Kultur wachsend. Infektion führt zu Schein-schwangerschaft. *C. psittaci*, Erreger der Papageienkrankheit.

### ***Rickettsia prowazekii* - 1.1 MB**

Erreger des epidemischen Typhus.  
Eng verwandt mit Mitochondrien. Enthält 24% 'tote Gene'

### ***Bordetella pertussis* - 3.88 MB**

Stäbchenbakterium. Erreger des Keuchhustens.

### ***Clostridium botulinum* - 4.01 MB**

Produziert Botulinustoxin, "Botulismus" Lähmungen.  
In verdorbenen Konserven

### ***Legionella pneumophila* - 4.10 MB**

Stäbchenbakterium. Erreger der Legionärskrankheit

### ***Neisseria gonorrhoeae* - 2.20 MB**

Fakultativ anaerobe Gonokokke. Erreger der Gonorrhoe (Tripper)

### ***Pseudomonas aeruginosa* - 5.90 MB**

Stäbchenbakterium, "Pfützenkeim".  
Verursacht Eiterungen (blaugrüner Eiter), Wundinfektionen

### ***Salmonella typhimurium* - 4.80 MB**

Obligat pathogenes Enterobakterium. Erreger des Typhus.

### ***Staphylococcus aureus* - 2.80 MB**

Kokke, produziert Enterotoxine.  
Verursacht oberflächliche und tiefe Eiterungen; Wundinfektionen

Abbildung 14: Die ersten pathogenen Organismen werden 1995 sequenziert.

Ein Sonderling ist auch *Synechocystis*, ein Cyanobakterium (früher als blau-grüne Alge bezeichnet), das zur Photosynthese befähigt ist und wahrscheinlich als Vorläufer der Chloroplasten aller grünen Pflanzen anzusehen ist.

1997 wurden, mit konventioneller Technik von Konsortien sequenziert, die Genome von *Escherichia coli* und *Bacillus subtilis* veröffentlicht. Während bei diesen Genomen eher das akademische Interesse überwog, sind die später etablierten Genomsequenzen (Übersichten in ref. [17, 18]) meist von unmittelbar medizinischer Relevanz.

*Helicobacter pylori* ist eine langsam wachsende, Flagellen tragende Spirochäte. Dank eines besonderen Enzyms, der Urease, kann sie im sauren Milieu der Magenschleimhaut leben, verursacht die häufigsten chronischen Infekte und ist bei 50% aller Menschen anzutreffen. Die Besiedelung der Magenschleimhaut ist häufig verbunden mit einer chronischen Gastritis und kann in schweren Fällen zu Magenkrebs oder Lymphomen führen. Der Ausgang hängt ganz vom bakteriellen Genotyp, von der Reaktion des ‚Wirtes‘ und von dessen Essgewohnheiten ab.

Mitte der 70er Jahre wurde erstmals in Connecticut das konzentrierte Auftreten einer ungewöhnlichen rheumatoiden Arthritis beobachtet, der *Lyme disease*. Zehn Jahre später fand man heraus, dass das Syndrom durch Zeckenbisse entsteht, bei denen eine neuartige Spirochäte, *Borrelia burgdorferi*,

übertragen wird. Merkwürdigerweise kann die Mikrobe nach der ersten Infektion monate- oder jahrelang im Wirt trotz einer heftigen Immunantwort überleben. Um diese ‚Resistenz‘ zu verstehen und womöglich einen Impfstoff zu finden, war es wichtig, zunächst das Genom zu entschlüsseln. Es zeigt gegenüber anderen Bakterien eine ungewöhnliche Struktur: es besteht aus einem linearen (!) Chromosom mit einer Größe von nur einer Megabase und mehreren kleineren zirkulären wie auch linearen Plasmiden, die sehr wahrscheinlich die Gene tragen, welche für die Virulenz verantwortlich sind.

Das Tuberkelbazillus (*Mycobacterium tuberculosis*) ist auch heute noch einer der am meisten gefürchteten Krankheitserreger. Nach dem Wiederaufleben von Infektionen sterben weltweit mehr als 3 Millionen Menschen pro Jahr an Tuberkulose. Immer häufiger werden gegen Antibiotika resistente Stämme registriert, so dass die Biologie dieses Mikroorganismus von großem Interesse ist. Die Sequenz des kompletten Genoms von *M. tuberculosis* deutet darauf hin, dass seine Widerstandsfähigkeit gegenüber Medikamenten wahrscheinlich genetisch determiniert ist. Ganz sicher wird die Kenntnis der zugrunde liegenden Resistenzmechanismen die Entwicklung neuer Therapiekonzepte befördern.

*Treponema pallidum*, der Erreger der Syphilis, besitzt eines der kleinsten prokaryotischen Genome, von nur wenig mehr als einer Million Basenpaaren. Es ist ein obligat humaner Parasit, der in künstlicher Umgebung kaum gedeiht. Deshalb weiß man wenig über seine Pathogenitätsmechanismen und besitzt noch keinen Impfstoff. Die DNA-Sequenz kann nun die notwendigen, aber auf andere Weise kaum erhältlichen, Informationen liefern, die man für biochemische Studien braucht.

Ein überall verbreiteter Mikroorganismus ist *Chlamydia trachomatis*. Es ist ein obligat intrazelluläres Pathogen und verursacht beim Menschen Entzündungen im Genitaltrakt und am Auge, die bis zur Erblindung führen können. Das Genom von *C. trachomatis* enthüllt ein eigenartiges Mosaik von Genen, die eigentlich eukaryotischen Ursprungs sind und auf eine Anpassung an die Lebensweise innerhalb menschlicher Zellen deuten.

Ähnlich kann sich auch *Rickettsia prowazekii*, der Erreger des epidemischen Typhus, nur in tierischen Zellen vermehren. Sein Genom hat überaus große Ähnlichkeit mit mitochondrialen Genomen, mehr als bei allen anderen bisher analysierten Mikroben. Dies ist vielleicht ein Hinweis, dass Rickettsien mit dem Vorläufer der Mitochondrien nahe verwandt sind. Es gilt inzwischen als gesichert, dass Mitochondrien in der Evolution von anderen Zellen in einer Art symbiotischen Prozesses aufgenommen wurden und so in der eukaryotischen Zelle die Rolle der Energieproduktion aus der ‚Verbrennung‘ von Nährstoffen mit Sauerstoff übernommen haben. Die Anpassung ursprünglich frei lebender Rickettsien an die Lebensweise in der Wirtszelle könnte ähnlich wie bei der Endosymbiose von eukaryotischen Zellen mit Mitochondrien verlaufen sein.

## 2.4 Genome von weiteren, genetisch relevanten Modellsystemen

Als besondere Leistung hervorzuheben ist die bereits 1998 von zwei Konsortien abgeschlossene Sequenzierung des Genoms des **Fadenwurms** (Nematoden) *Caenorhabditis elegans*, denn mit diesem Genom wurde das bis dahin größte eines eukaryotischen Organismus entschlüsselt [19]. Die 97 Millionen Basenpaare von *C. elegans* enthalten etwa 19.000 Gene. Das Besondere an diesem Organismus ist, dass er aus etwa 1000 Zellen besteht, von denen 302 Nervenzellen sind. Durch neurobiologische Studien, die bereits in den 70er Jahren begannen, kennt man nicht nur die Lokalisation und den Ursprung der Entwicklung jeder einzelnen Nervenzelle, sondern auch die synaptischen Verschaltungen zwischen den Neuronen [20]. Diese Ergebnisse konnten nun mit den Genen korreliert werden, die für Gehirnfunktionen verantwortlich sind. Da die meisten molekularen Komponenten des Vertebratengehirns bereits bei diesem primitiven Vorläufer vorhanden sind, liefert *C. elegans* sozusagen ein Minimalmodell für das Nervensystem aller Tiere.

Seit dem Jahre 2000 ist die komplette Sequenz des **Genoms von *Drosophila*** verfügbar [21]. Die vier Chromosomen umfassen 180 Mio. Basenpaare mit etwa 20.000 Genen. Das Unternehmen wurde in Zusammenarbeit von *Celera Genomics* und mehreren, vom *National Institute of Health* (NIH) finanzierten, Genzentren durchgeführt. Das akademische Team stellte eine physikalische Genkarte zur Verfügung, die Firma produzierte die meisten Sequenzdaten. Bemerkenswert ist, dass *Celera* die Sequenzierung nach der ‚shotgun‘-Methode durchführte und es gelang, die Sequenzen weitgehend durch ausgeklügelte Computerprogramme zusammenzusetzen.

## 2.5 Das Humangenomprojekt

### 2.5.1 Vorgeschichte und Konzept

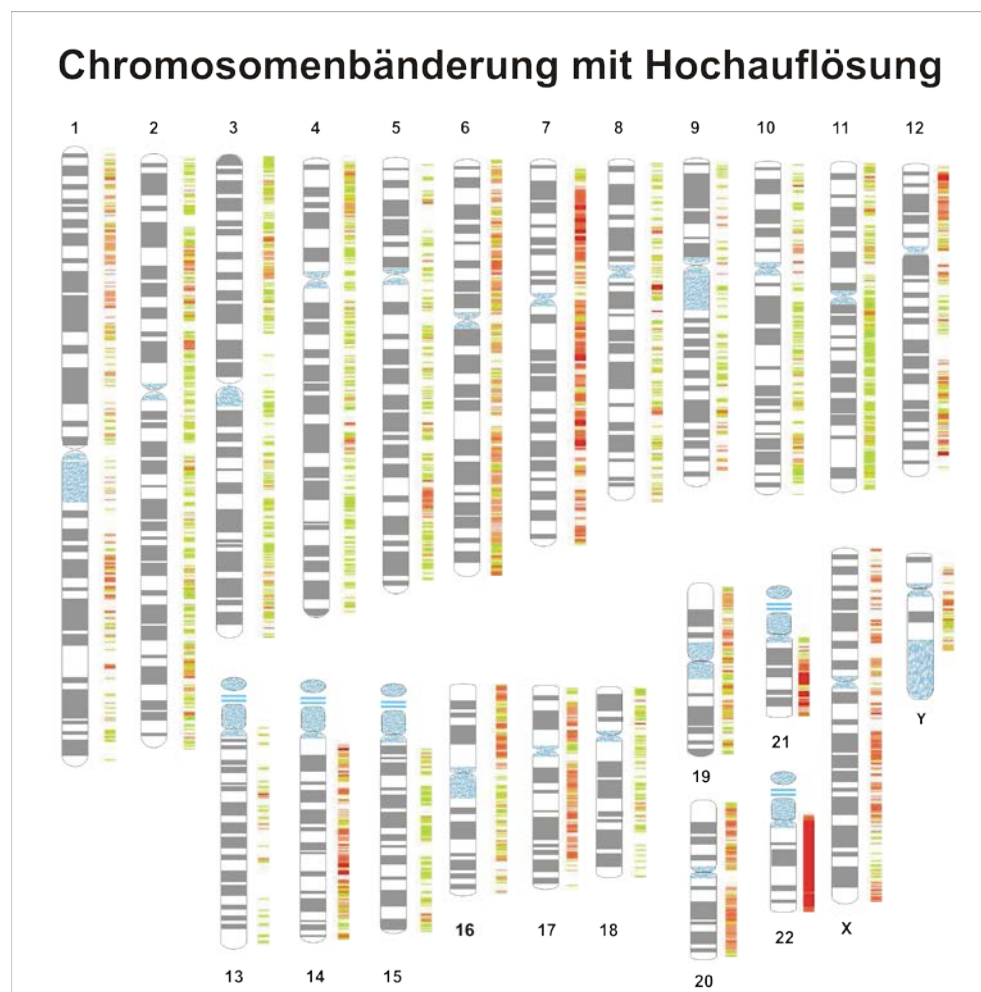
Das Konzept für das **Humangenomprojekt** reicht in die Mitte der 80er Jahre zurück und wurde von den Forschern und der Öffentlichkeit in aller Breite bis zum Ende des Jahrzehnts erörtert. In den USA wurden schließlich vom Kongress das *Department of Energy* (DOE) und das *National Institute of Health* (NIH) für die verantwortliche Planung und Durchführung des Projektes eingesetzt [22]. 1988 wurde von diesen beiden Institutionen ein gemeinsames Memorandum verabschiedet und ein aus drei gleich langen Abschnitten bestehender Fünfzehnjahresplan aufgestellt, dessen Beginn auf das Jahr 1990 festgelegt wurde, und für dessen Durchführung Gesamtmittel in Höhe von 3 Milliarden US Dollar veranschlagt waren. Hauptziel des wohl ehrgeizigsten Forschungsvorhabens der Genomforschung war die vollständige Sequenzierung des menschlichen Genoms. Dazu wurden eigens einige große Forschungszentren etabliert, aber auch kleinere Laboratorien, die sich speziellen Fragen widmeten, finanziell unterstützt. Die 1988 als Dachorganisation gegründete *Human Genome Organisation* (HUGO) hat eine weltweite Beteiligung am Humangenomprojekt initiiert [23]. Heute führt HUGO eine vollständige Dokumentation über Humangensequenzen und genetische Krankheiten [24].

War man bisher - wie auch auf anderen Feldern der Molekularbiologie - vom Erscheinungsbild, oder günstigstenfalls von der Funktion eines Proteins, ausgegangen, um das entsprechende Gen zu



isolieren, so wollte man nun auch hier den umgekehrten Weg beschreiten. Insbesondere die Suche nach Genen, welche für menschliche Erbkrankheiten verantwortlich sind, wurde immer schon intensiv betrieben, es war aber meist überaus arbeits- und zeitaufwendig, das für eine bestimmte Erbkrankheit verantwortliche Gen zu identifizieren und darüber hinaus auch seine Funktion zu beschreiben. Kannte man nur das Erscheinungsbild einer Erkrankung, so musste zunächst durch die genetische Analyse von allen Mitgliedern einer betroffenen Großfamilie der Genort eingegrenzt werden, um dann die entsprechende Region durch Klone abzudecken und ‚Kandidatengene‘ durch Sequenzieren zu identifizieren. Beispielsweise hat es 15 Jahre in Anspruch genommen, das für die Duchenne'sche Muskeldystrophie verantwortliche Gen zu finden und seine Funktion zu ermitteln.

Vorgesehen war, die Sequenz aller 3,5 Milliarden Basenpaare des menschlichen Genoms (also aller 24 Chromosomen; Abbildung 15) bis zum Jahre 2005 entschlüsselt und alle Gene identifiziert zu haben. Ohne Kenntnis der genomischen Sequenz war man vorläufig auf Schätzungen angewiesen, welche zunächst auf 80.000 bis 100.000 menschliche Gene deuteten. Nachdem das Genom nun entschlüsselt ist, rechnet man nur noch mit rund 35.000. Die Unsicherheit in der Abschätzung der Zahl von Genen bei höheren Organismen ist darin begründet, dass nicht nur ihre Genstrukturen sondern auch die Strukturen der Genome wesentlich komplexer sind als bei niederen Organismen, so dass Extrapolationen nicht möglich sind.



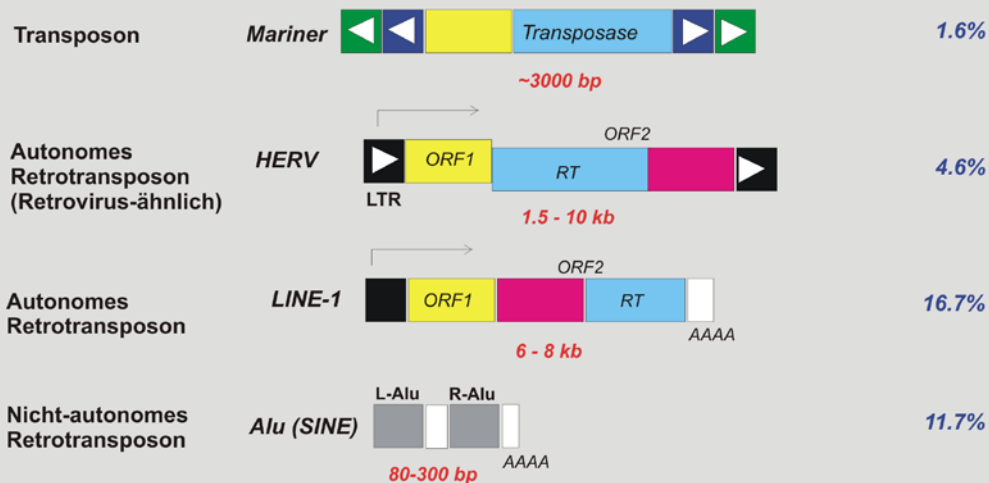
*Abbildung 15: Der menschliche (haploide) Chromosomensatz. Heterochromatische Regionen sind blau dargestellt, die Bänderung (grau) des Euchromatins ergibt sich aus verschiedenen Färbemethoden. Autosomen: 1 – 22; Geschlechtschromosomen X und Y. Jeweils rechts daneben eine Feinanalyse der Bänderung.*

Bei Mikroorganismen sind die Gene sehr dicht gepackt. Bei Bakterien schließt ein Gen fast an das andere an, manchmal werden sogar beide DNA-Stränge zur Information benutzt. Auch in dem niederen Eukaryoten Hefe werden 75% der gesamten DNA-Sequenz zur Codierung von Proteinen genutzt, im Mittel beansprucht ein Gen also etwa 2 Kb. Selbst beim Fadenwurm *C. elegans* ist die Gendichte nicht viel geringer, hier beansprucht ein Gen im Mittel 5 Kb. Würde man mit diesem Wert auf das Humangenom hochrechnen, so müsste der Mensch (vergleichbar auch andere Säugetiere) etwa 600.000 Gene besitzen. Wie kommt also der reale Wert von weniger als 20.000 menschlichen Genen zustande?

Aus verschiedenen Untersuchungen weiß man schon lange, dass in der menschlichen DNA die für Proteine codierenden Sequenzen in der Tat nur etwa 2-3 % ausmachen. Ein weitaus größerer Prozentsatz entfällt auf die schon erwähnten Introns. Während Introns bei niederen Eukaryoten eher selten vorkommen und auch nicht besonders groß sind, kann ein Säugergen durch mehr als 60 Introns unterbrochen sein, die noch dazu jedes mehrere KB, unter Umständen sogar über 100 KB, lang sein können! Eine tierische Zelle benötigt also weitaus mehr DNA pro Gen als ein niederer Organismus.

Den Löwenanteil in der menschlichen DNA machen mit mindestens 45% die ‚repetitiven Sequenz-Elemente‘ aus. Verschiedene Typen dieser Elemente (mit einer Länge zwischen 300 und 8.000 Basenpaaren) haben sich in hoher Kopienzahl über das menschliche Genom verbreitet, weil sie grundsätzlich in der Lage sind, neue Kopien von sich selbst zu erzeugen, die in andere Positionen im Genom springen können. Das Vorkommen solcher ‚mobilen‘ Elemente im Humangenom ist nicht überraschend, hat man doch gleichartige Elemente in allen bisher untersuchten Organismen gefunden. Wenn auch ihre Herkunft im Dunkel liegt, so belegt die heutige Struktur der größten Gruppen dieser Elemente doch eine enge Verwandtschaft zu den so genannten Retroviren, die sich als ‚genetische Schmarotzer‘ in ein Wirtsgenom einnisten können. Während die Mehrzahl der Elemente im menschlichen Genom heute vermutlich stillgelegt ist und glücklicherweise einen Dornröschenschlaf hält, findet man gelegentlich doch Patienten, bei denen ein solches Element in ein lebenswichtiges Gen gesprungen ist und dadurch dessen Funktion eliminiert hat. Nicht zu unterschätzen ist im menschlichen Genom die Zahl so genannter Pseudogene. Sie sind daran erkennbar, dass sie keine funktionellen Leseraster besitzen. Der größte Teil der Pseudogene geht wahrscheinlich auf alte Genduplikationen zurück, bei denen die Evolution gewissermaßen in einer Sackgasse endete (Abbildung 16).

## Repetitive Elemente im Humangenom



Jede Art von diesen (und weiteren) Elementen können im Genom "springen". Ihre Anwesenheit verursacht Störungen in der Ablesung der betroffenen Gene. Die Muster können vererbt werden und beim Träger schwere Krankheiten auslösen.

**Heute wird der Anteil solcher Elemente im Humangenom auf mindestens 45% geschätzt.**

Krankheiten des Menschen, die durch Transposons verursacht werden

Gene, die durch Insertionen von **Alu-Elementen** unterbrochen sind:

- Factor IX Gen (Haemophilie B)
- NF1 Gen (Neurofibromatose)
- FGFR2 Gen (Apert Syndrom)
- APC Gen (Dickdarm-Krebs)
- XLA und XSCID Gene (X-gekoppelte Immundefizienz)
- BRCA2 Gen (Brust-Krebs)

Gene, die durch verstümmelte Kopien von **LINE-1** unterbrochen sind (in der Keimbahn oder in der frühen Entwicklung):

- Factor VIII Gen (Haemophilie A)
- DMD Gen (Duchenne'sche Muskeldystrophie)
- APC Gen (Dickdarm-Krebs)
- β-Globin Gen

Abbildung 16: „Überflüssige Elemente“ im Humangenom?

### 2.5.2 Strategien und Erfolge

Wenn man sich in unbekanntem Gelände bewegen will, benötigt man gute Karten. So sollte in den ersten Jahren des Humangenomprojektes eine genaue genetische Karte aufgestellt werden, die möglichst viele nahe bei einander liegende Marker (also bekannte Sequenzen) verzeichnete, sowie physikalische Genkarten, welche gestatteten, das gesamte Genom mit möglichst langen DNA-Abschnitten aus Klonbibliotheken geordnet und lückenlos zu überdecken. Im Sinne einer 'gerichteten Strategie' sollte dann die Sequenzierung der Klone mit schnellem und kostengünstigem Durchsatz folgen. Außerdem mussten die notwendigen Grundlagen für die Verarbeitung und Analyse großer Datenmengen geschaffen werden.

Als sich 1993 herausstellte, dass der erste Fünfjahresplan praktisch bereits erfüllt war, wurden sofort die Ziele eines zweiten Fünfjahresplanes für 1994 -1998 festgelegt. Auch dieses Programm erwies sich als erfolgreich. 1998 wurde dann schließlich der dritte Fünfjahresplan angegangen, welcher nun außer der vollständigen Sequenzierung des Humangenoms auch dessen funktionelle Analyse und die Bestimmung der genetischen Varianzbreite beinhaltete [25].

Ein erster Meilenstein im Humangenomprojekt wurde Mitte 1999 mit der Aufklärung der Sequenz von Chromosom 22 erreicht [26]. Zum ersten Mal konnten wir uns ein Bild machen, wie ein menschliches Chromosom strukturiert ist. Beteiligt an der Entzifferung der 33,5 Millionen Bausteine nach der ‚Klon-für-Klon-Strategie‘ waren Forscher aus dem *Sanger Centre* (Cambridge), den Universitäten von Oklahoma und Washington sowie der Keio-Universität in Japan. Weil auf dem kurzen Arm von Chromosom 22 praktisch keine für Proteine codierenden Gene vorkommen, sondern fast ausschließlich die kaum sequenzierbaren repetitiven Sequenzen, konzentrierten sich die Forscher auf den langen Arm, der übrigens genreicher zu sein scheint als andere Chromosomen. Bald folgte die Sequenz von Chromosom 21 [27], an welcher auch deutsche Forscher mitgearbeitet haben.

Im September 1999 kamen die führenden Wissenschaftler des Humangenomprojektes überein, bereits im Sommer 2000 eine Art ‚Rohfassung‘ (*working draft sequence*) von mindestens 90% des Humangenoms bereitzustellen. Das Erreichen dieser Absicht erschien greifbar, nachdem eine neue Generation von superschnellen Sequenzierautomaten zur Verfügung stand. Eine solche Maschine kann mehrere hundert Kilobasen an Rohdaten am Tag liefern, und in den Zentren wurden sie in großer Zahl parallel eingesetzt. Auf diese Weise lagen zu Beginn 2000 etwa 17% (~535 Mb) der Sequenz des Humangenoms in abgesicherter Form und etwa 47% (1,524 Mrd. Bp) in vorläufiger Form vor.

Die Beschleunigung des Humangenomprojektes geht ohne Zweifel auch auf die Initiative *Craig Venter*, des damaligen Direktors von *Celera Genomics*, im Jahr 1998 zurück, das Humangenom kurzerhand mit der *shotgun*-Technik anzugehen. Zwar wurde der ursprüngliche Plan, eine komplette und hoch gesicherte eigene Sequenz (mit einer zehnfachen Abdeckung) zu produzieren aufgegeben, dafür aber die Absicht forciert, die von *Celera* in rascher Folge gesammelten Daten (mit einer etwa fünffachen Abdeckung) möglichst bald mit den öffentlich zugänglichen Daten aus dem Humangenomprojekt zu vereinen. In der Tat wurde im Januar 2000 berichtet, *Celera* habe bereits 81% der Sequenz des Humangenoms eingefahren und könne zusammen mit den öffentlichen Daten nun 90% des Genoms überdecken. Zu Recht ist über dieses Vorgehen ein heftiger Disput entstanden, denn die Teilnehmer am Humangenomprojekt hatten sich verpflichtet, ihre Daten jeweils innerhalb von 24 Stunden publik zu machen. Sie wehrten sich vor allem dagegen, dass *Celera* öffentliche Daten zur Aufbesserung seiner eigenen benutzen und alle Daten zusammen dann Interessenten gegen Lizenzgebühren zur Verfügung stellen wollte. Anfang April 2000 schlug dann die Nachricht wie ein Blitz ein, *Celera* habe 99% des Humangenoms fertig gestellt. Allerdings mussten die Sequenzen noch in der richtigen Reihenfolge angeordnet werden, eine nicht ganz leichte Aufgabe, wenn man bedenkt, dass *shotgun*-Sequenzen nur mit Hilfe aufwendiger Computerprogramme zusammengesetzt werden können und die vielen Repetitionen im Humangenom eine weitere Komplikation bedeuteten. Die

Spezialisten von *Celera* verwiesen jedoch optimistisch auf das Genom von *Drosophila*, welches sie mit der gleichen Strategie kurz zuvor bewältigt hatten.

Trotz aller Differenzen einigten sich beide Parteien im Sinne des ursprünglichen Zieles. Ende Juni des Jahres 2000 gingen sie (sozusagen Hand in Hand mit dem amerikanischen Präsidenten) mit der bedeutsamen Meldung an die Öffentlichkeit, dass eine vorläufige Fassung der gesamten Sequenz des Humangenoms zur Verfügung stehe, und dass sie gemeinsam an der Vervollkommnung der Daten weiter arbeiten würden. Mitte des Jahres 2001 kamen dann die getrennten Publikationen heraus [28, 29]. So wird wohl die Gefahr vermieden, dass das Humangenom einseitig für kommerzielle Interessen mit Beschlag belegt wird. Allen Beteiligten wurde klar, dass nur eine konstruktive Zusammenarbeit die noch vor ihnen liegenden Aufgaben bewältigen konnte. Noch lange aber schwelte der Streit, ob *Celera* tatsächlich in der Lage gewesen wäre, ohne die Zuhilfenahme der Daten aus dem Humangenomprojekt seine Fassung zu präsentieren [30]. Als erstes hat man schon anhand der Rohfassung begonnen, in den assemblierten Sequenzen die Gene aufzuspüren und ihre Funktionen so weit wie möglich zu beschreiben. Bis zum Jahre 2003 sollte dann die vollständige und durch weitere Arbeiten gesicherte genomische Sequenz fertig gestellt sein. Hierzu sind verschiedene Programme aufgelegt worden, unter anderem solche, die nochmals das gesamte Genom in Form überlappender Klone von jeweils mehreren hundert Kilobasen darstellen und auch geeignet sind, die genetische Karte zu verfeinern, oder solche die erlauben, alle gefundenen Gene auf den Chromosomen gewissermaßen als einzelne farbige Punkte abzubilden (Abbildung 17), bzw. ganze Chromosomen „anzufärben“ (Abbildung 18).

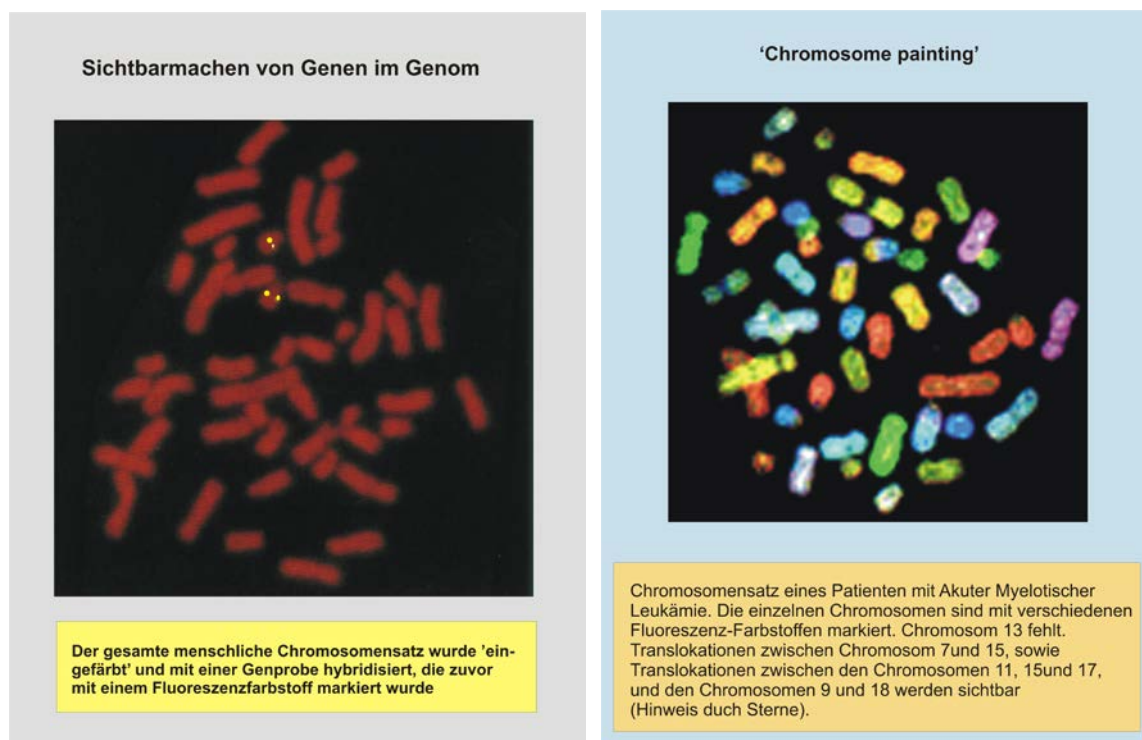


Abbildung 17: Markierung von Genen auf Chromosomen.

Abbildung 18: Chromosome painting zur Feststellung von Translokationen.



Die endgültige Sequenz des Humangenoms ist seit 2003 bzw. 2006 verfügbar [31].

### 2.5.3 Neue Aspekte im Humangenom-Programm

Einen vorrangigen Platz sollte die Identifizierung menschlicher Krankheitsgene einnehmen (Abbildung 19). Von den vermuteten 3000 Genen für so genannte monogene Erbkrankheiten konnte bisher erst etwa die Hälfte ermittelt werden, obgleich in den Jahren seit Beginn des Humangenomprojektes die Rate der Ermittlung steil angestiegen ist.

## Genloci auf den menschlichen Chromosomen

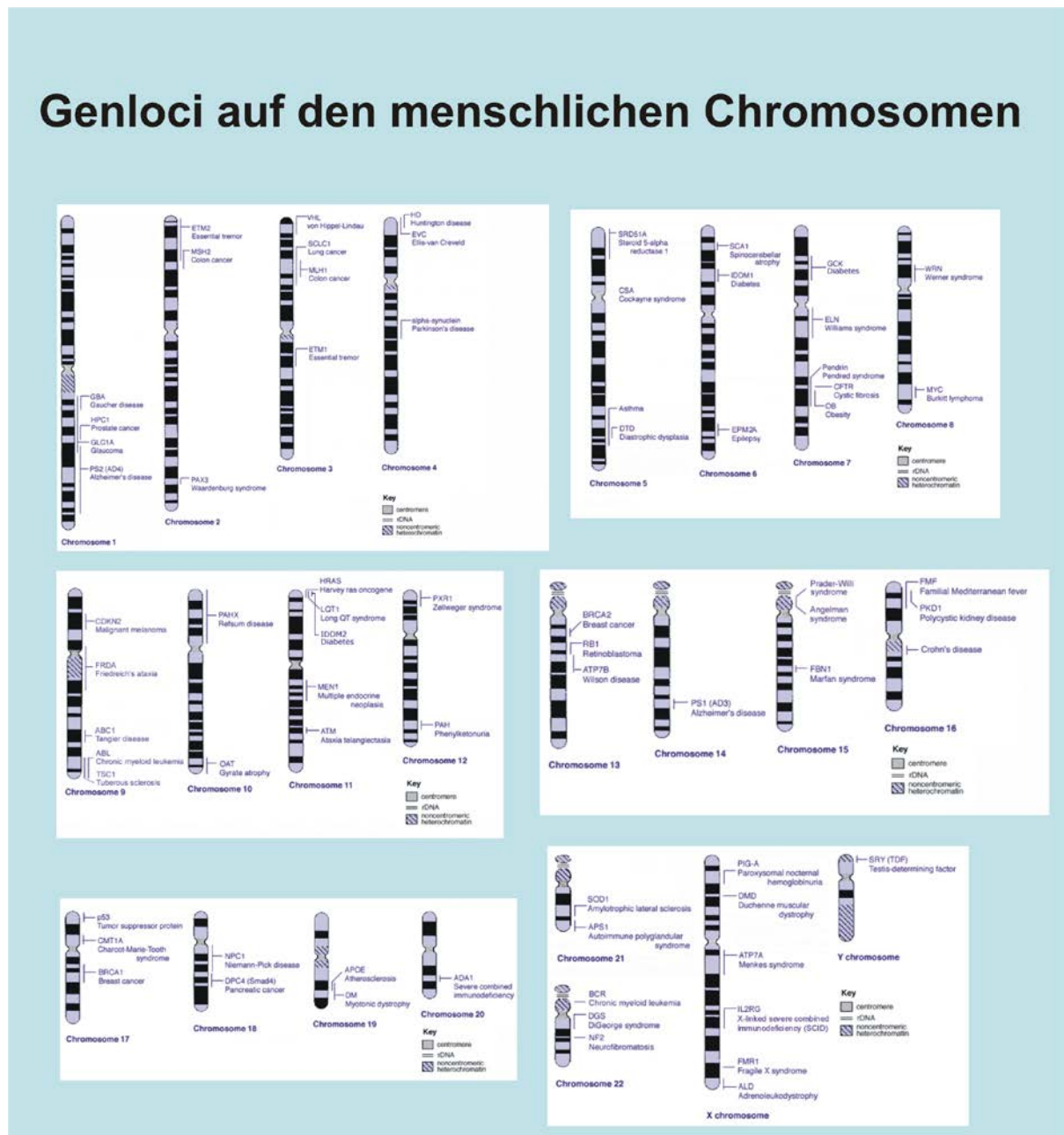


Abbildung 19: Auf menschlichen Chromosomen lokalisierte Defekte bei genetischen Krankheiten (englische Beschriftung).

Schwieriger dagegen bleibt die Identifizierung von Genen, die zur Entstehung multigener Krankheiten, wie Diabetes, Koronarkrankheiten, neurodegenerativen Erkrankungen oder geistigen Anomalien beitragen. Wachsende Bedeutung wird die Analyse der ‚Krebsgene‘ haben, von denen mittlerweile weit über 100 bekannt sind. Das *Cancer Genome Anatomy Project* wurde 1999 als interdisziplinäres Programm mit dem Ziel gestartet, alle Informationen zur Aufklärung der molekularen Ursachen der Krebsentstehung zu gewinnen. Dazu hat man laufend sog. ESTs („Expressed Sequence Tags“) gesammelt, das sind 400-500 lange Sequenzen, die aus der mRNA der involvierten Gene durch Zurückschreiben in DNA-Fragmente generiert werden. Heute umfasst die in Datenbanken niedergelegte Information etwa 1,5 Mio solcher EST-Sequenzen [32].

Ein weiterer Aspekt des Humangenomprojektes war es, die Variabilität des menschlichen Genoms zu untersuchen (HGDP, *Human Genome Diversity Project*) [33]. Man wusste bereits 1997, dass individuelle Unterschiede in einzelnen Basenpaaren (SNP= *single nucleotide polymorphisms*) relativ häufig auftreten, wahrscheinlich an jeder 1.000sten Stelle im Genom. Ein großer Teil dieser Varianten ist dafür verantwortlich, dass Krankheiten individuell verschieden ausgelöst oder unterschiedliche Reaktionen auf Pharmaka beobachtet werden. So haben zehn führende pharmazeutische Firmen 45 Mio US\$ für die Kartierung von Krankheit auslösenden Mutationen beim Menschen bereitgestellt. Die Daten sind größtenteils bereits verfügbar und frei zugänglich. Man hofft, durch die Ergebnisse individuellen Medikationen näher zu kommen und dadurch Nebeneffekte von Medikamenten zu reduzieren. Außerdem können nun Krankheitsgene schneller als bisher identifiziert werden. Vorgesehen ist zum Beispiel ein Vergleich der Gene von Patienten mit Koronarkrankheiten mit denen gesunder Probanden.

Im Jahr 2003 sollte auch die gesamte Sequenz des Mausgenoms zur Verfügung stehen. Die Maus gewann als ‚Modellsystem‘ für menschliche Erkrankungen in allen Bereichen der Biomedizin an Bedeutung, stimmen doch die Funktionen 98% all ihrer Gene mit denen des Menschen überein. Schon jetzt sind umfangreiche Studien im Gange, um die Folgen der gezielten Ausschaltung, der Modifizierung oder Fehlaktivität von Mausgenen zu verstehen, denn aus begreiflichen Gründen sind solche wichtigen Funktionsanalysen nicht am Menschen direkt durchführbar.

## 2.6 Andere Genomprojekte im Fortschritt

### 2.6.1 Menschliche Parasiten

Die folgenden Jahre haben uns weitere interessante genomische Sequenzen beschert. Die Sequenzierung der Genome dreier bedeutsamer **menschlicher Parasiten**, *Plasmodium falciparum* [34], dem durch Anopheles-Mücken [35] übertragenen Erreger der Malaria [36]; *Trypanosoma brucei* [37], dem durch Tsetse-Fliegen übertragenen Erreger der Schlafkrankheit [38]; und *Leishmania major* [39], dem Verursacher der Leishmaniose (Orientbeule) [40], wurde intensiv betrieben und bis spätestens 2007 abgeschlossen.

### 2.6.2 Nutzpflanzen

Auch das bis zur Jahrtausendwende bestehende Defizit in der **pflanzlichen Genomforschung**, das in erster Linie in der enormen Genomgröße der meisten Pflanzen zu suchen ist, wurde intensiv angegangen. Bereits 2000 wurde das Genom von *Arabidopsis thaliana*, einem Verwandten unserer vielen Kohllarten, der wegen seines ungewöhnlich kleinen Genoms von 130 Megabasen (mit ca. 20.000 Genen) als dicotyle Modellpflanze fungierte, entschlüsselt [41]. Reis, von dem mehr als die Hälfte der Weltbevölkerung zehrt, stellte den zweiten, weltweit bearbeiteten pflanzlichen Modellorganismus dar. Die ‚nur‘ 430 Millionen Basenpaare der Genome zweier Reissorten wurden durch internationale Consortien (Japan, Großbritannien, Korea, Frankreich, China und USA) bearbeitet [42], bis heute scheint es aber nur vorläufige Gesamtsequenzen („draft sequences“) zu geben [43].

Von anderen wichtigen Nutzpflanzen, wie Mais, Raps, Sojabohne, Baumwolle, Kartoffel oder Gerste, deren Genome ebenfalls zum Teil größer als das Humangenom sind, wurden zuerst nur vergleichende Sequenzierungen unternommen. Das Schwergewicht lag dabei auf der Erkennung oder Erzeugung genetischer Diversität, denn gerade Pflanzen weisen als Folge der an sehr unterschiedliche Standorte über lange Zeiträume adaptierten Genmuster eine außerordentlich große genetische Vielfalt auf. Diese zu nutzen ist von mindestens so großer Bedeutung wie der Anbau transgener Kulturpflanzen. Heute sind fast Tausend pflanzliche Genome bekannt [44]. Dazu gehören auch die von diversen Algen, Moosen, Farnen, oder einkeimblättrigen (monocotylen) Pflanzen, vor allem aber die der gängigsten Kulturpflanzen. Allein 14 zu den Gräsern gehörende Getreidearten sind sequenziert worden; das Genom von Mais bricht den Größenrekord mit 37 Billionen Basenpaaren (also tausendmal mehr als das menschliche Genome mit 3,5 Milliarden). Die Genome aller üblichen Gemüse-, Frucht- und Zierpflanzen sind erhältlich. Der Leser mag selber den Katalog [44] durchgehen, um die vielen Arten und den Status der Projekte zu sehen.

### 2.6.3 Tiere

Der ‚Zoo‘ der von Tieren bis heute erstellten Genomen beginnt bei den ‚niederen‘ Lebewesen: Schwämmen, Medusen, Muscheln (Auster), Seeigeln, Lanzettfischchen, und setzt sich mit Fröschen und Knochenfischen (8 Arten) fort. Unter Wirbeltieren finden sich die Genome von sechs Reptilien und acht Vögeln.

Bei den Insekten sind es allein 21 *Drosophila*-Arten, deren Genomsequenzen verglichen wurden. Auch die Genome nützlicher und schädlicher Insekten befinden sich im Repertoire: Bienen, Wespen, Fliegen, Laus, Motte, Wasserfloh und Schmetterlinge.

Bei den Säugetieren liest sich das Register wie eine Kurzfassung von Brehms Tierleben, jeweils mehrere Vertreter wildlebender Tiere (sogar ausgestorbener Arten), viele Zootiere sowie alle Nutz- und Haustiere sind im Katalog vertreten [45].

#### 2.6.4 Mikroben

In den ersten Jahren seit dem Aufkommen von Genomprojekten hat auch ein regelrechter ‚Run‘ auf die Sequenzierung **mikrobieller Genome** eingesetzt. Die Liste der genomischen Sequenzen, die bisher von bakteriellen Krankheitserregern ermittelt wurden, liest sich wie das Verzeichnis eines Handbuchs der Infektionskrankheiten. Zusätzlich zu den ‚öffentlichen‘ gibt es eine Vielzahl von Projekten, die von Pharmafirmen durchgeführt werden und deren Daten bis heute zum Teil nicht frei zugänglich sind - zumindest solange, bis aussichtsreiche Anwendungen daraus patentiert werden können. Der Spielraum bei den Mikroorganismen scheint unermesslich groß, schätzt man doch, dass nur etwa 10% der auf der Erde lebenden überhaupt bekannt sind. Vergessen sollte man unter den Krankheitserregern auch nicht verschiedene Arten von ‚**Protisten**‘ (Dinoflagellaten, Amöben, und andere pathogene Kleinlebewesen), an denen großes Interesse besteht, weil sie beim Menschen und bei Tieren schwerwiegende Infektionen auslösen [45].

Besonders attraktiv wurden für die Zukunft Mikroorganismen mit ungewöhnlichen Stoffwechselleistungen, die beispielsweise zur Synthese neuer Biomoleküle, zur Entwicklung neuartiger Bioprozesse, wie Rohstoffrückgewinnung, Entsorgung von Abwässern, Bereitstellung umweltfreundlicher Energien, oder zur Beseitigung oder Umwandlung bisher nicht ‚verdaubarer‘ Rückstände eingesetzt werden können (vgl. auch Abschnitt 3.1.3).

Um nur eines der wichtigen Ergebnisse zu nennen: Bereits 1997 wurde das Genom des Archebakteriums *Methanobacterium thermoautotrophicum* entschlüsselt [46], das von technischem Interesse ist, und zugleich ein Beispiel dafür bietet, wie man die aus Vergleichen mit anderen Genomen gewonnene Information zu bekannten (oder wahrscheinlichen) Genfunktionen darstellen kann. Die Archeen werden seit langem als eigenes ‚Reich der Lebewesen‘ geführt; einen Überblick findet sich in ref. [45]. Inzwischen sind die Genome von über 30 verschiedenen Spezies von Methanobakterien und anderen nützlichen Archebakterien bekannt geworden. Die Organismen werden industriell zur Klärung von Abwässern oder zur Biogas-Produktion eingesetzt.

#### 2.6.5 Begonnene und laufende Genomprojekte

Umfangreiche Genomprojekte brauchen eine kostspielige Förderung, und so beteiligen sich daran (z.B. in den USA) viele renommierte Forschungsinstitute, staatliche Geldgeber und private Stiftungen gemeinsam an diesen Unternehmungen. Im Folgenden wird auf eine Liste von Projekten Bezug genommen, an denen das *J. Craig-Institut* beteiligt ist [47].

##### 2.6.5.1 Künstliche Bakterienzellen

Im Jahre 2003 war es gelungen, ein kleines, Bakterien infizierendes Virus auf synthetischem Wege zu schaffen. 2008 konnte man ein kleines bakterielles Genom (*Mycoplasma mycoides*) synthetisieren,

das aber nicht aktiviert werden konnte. Erst 2010 gelang der Durchbruch: Die etwa eine Mio Basenpaare große und vollständig chemisch erzeugte DNA wurde in einer Hefezelle zusammengesetzt, und in dieser entwickelten sich schließlich Bakterien, die sich selbst vermehren konnten.

#### **2.6.5.2 ‚Synthetic Genomics‘**

Die Schaffung synthetischen Genmaterials, welche chemische Methoden und Computerprogramme benutzt, wird Biologen und Ingenieuren gestatten, solch genetisches Material zu erstellen, das nur schwer zugänglich oder nicht verfügbar ist: Man wird Gene, Chromosomen oder gar Genome schaffen, um neue Therapeutika oder Biotreibstoffe biotechnologisch zu produzieren oder schnellstmöglich Impfstoffe gegen neu auftkommende Erreger zu entwickeln. Selbstverständlich sind Bestrebungen im Gange, Chancen und Risiken abzuwägen, in Verordnungen umzusetzen, auch unter Berücksichtigung sicherheits-technischer, ethischer, soziologischer und ökonomischer Gesichtspunkte.

#### **2.6.5.3 Human-Mikrobiom-Projekt**

Das ‚human microbiome projekt‘ (HMP) wurde durch das *National Institute of Health* (NIH) im Jahre 2012 gestartet [48]. Es soll die Analyse aller Lebewesen umfassen, die in und auf dem menschlichen Körper leben, schätzungsweise sind das etwa  $10^{14}$ , 10-mal mehr als der menschliche Körper Zellen besitzt. Unterprojekte kümmern sich um die Hautoberfläche, Körpersekrete, Körperöffnungen, innere Organe, u.dgl. Das Ziel ist, Wechselbeziehungen zu menschlichen Krankheiten oder Unwohlsein zu finden; geplant ist eine Datenbank, die etwa 1000 neue Mikroorganismen beschreibt. Man möchte auch untersuchen, ob alle Arten von Mikroben über den ganzen Körper verteilt sind, oder sich bestimmte Gruppen in begrenzten Körperbereichen aufhalten. Viele medizinische Kliniken beteiligen sich mit speziellen Fragestellungen an diesem Projekt.

#### **2.6.5.4 Erforschung aller Kleinlebewesen in den Weltmeeren**

Seit 2003 engagieren sich Forscher aus mehreren Ländern (USA, England, Schweden, Finnland, Norwegen, Dänemark, Estland und anderen baltischen Staaten), die Kleinlebewesen in allen Meeren zu sammeln, zu analysieren und deren Genome zu sequenzieren [49]. Während wir viele erd- oder luftgebundene Mikroben kennen, bleibt für uns die Welt der entsprechenden Meeresbewohner noch verschlossen. In einer ersten Expedition (2003 bis 2008) sammelten die Forscher an Bord eines Forschungsschiffes Proben in beiden Küstenregionen von Nordamerika, durchforsteten beide großen Ozeane, Atlantik und Pazifik, und wagten sich nach Australien sogar in die Antarktis. Eine zweite Expedition (2009 und 2010) war der Erforschung der Karibik, des Nordatlantiks und der von den Weltmeeren abgeschnittenen ‚europäischen‘ Meere (Ostsee, Mittelmeer und Schwarzes Meer) gewidmet. Es wurden systematisch Proben in unterschiedlichen Wassertiefen genommen, durch



immer feiner werdende Filter gesiebt, und alle Proben schließlich in die Labors geschickt, um die Arten dieser Kleinlebewesen zu bestimmen und ihre DNA zu analysieren.

#### **2.6.5.5 Das erste diploide Genom eines einzelnen Individuums**

C. Venters genomische DNA war die erste, die bestehend aus den Chromosomensätzen beider Elternteile, mit 20 Mrd. Basenpaaren im Jahre 2007 vollständig angegeben werden konnte [50]. Im Vergleich zu den Sequenzen des ‚Humangenomprojektes‘ kamen bei dem individuellen Genom etwa siebenmal mehr genetische Variationen von Mensch zu Mensch zum Vorschein, insgesamt etwa 4,1 Mio Varianten. Die Variationen in der Sequenz nehmen 12,3 Mio Basenpaare ein, davon 3,2 Mio sogenannte ‚single nucleotide polymorphisms (SNPs)‘, 1,2 Mio vorher nie beobachtete Variationen, und fast 1 Mio nicht-SNP Variationen. Das ist nur der Anfang: Heute können individuelle Genome kommerziell zu günstigen Preisen (< 10.000 US\$) sequenziert werden.

#### **2.6.5.6 Südafrika Genomprojekt**

Ein weltweit anerkanntes Projekt begann unter Führung der ‚Human Genomic Medicine Group‘ des JCVI, das 2010 veröffentlicht wurde. Es sollte die verdeckten genomischen Variationen zwischen verschiedenen negroiden Stämmen untersuchen, um ihre Bedeutung für Krankheitsrisiken und in der medizinischen Behandlung zu verstehen. Herangezogen wurden zunächst drei Buschmänner und ein Bantu (letzterer nicht zufälligerweise der Erzbischof Desmond Tutu). Man fand 1,3 Mio Varianten, die bisher nirgendwo aufgetaucht waren. Jedenfalls ergab sich, dass Südafrikaner genetisch anders sind als Europäer, Asiaten oder sogar Westafrikaner. Die Analyse gab auch Einblick in den (genetisch bedingten) Alterungsprozess, denn die Probanden waren alle über 70 Jahre alt.

### 3 Gene als ‚Material‘

Die neu entstandene Gentechnik hatte sehr bald Erfolge vorzuweisen. Zwischen 1977 und 1980 waren viele verschiedene Gene aus unterschiedlichen Organismen, wie Bakterien, Hefe und tierischen Zellen kloniert und sequenziert worden. Eine besonders wichtige Entdeckung war, dass die meisten eukaryotischen Gene mosaikartig aufgebaut sind: Codierende Gensegmente werden durch (oft sehr lange) nicht-codierende Segmente (so genannte Introns) unterbrochen, die zur Erstellung der Boten-RNA entfernt werden müssen. Dies wurde sehr früh schon an der Struktur der Globingene und von Immunglobulingenen des Menschen und der Maus deutlich. Genstrukturen zu analysieren, war aber nicht das alleinige Ziel.

#### 3.1 'Manipulierte' Gene - Chancen und Risiken

##### 3.1.1 Kategorien der Biotechnik

Intensive Bemühungen galten der Frage, wie man klonierte Gene zur Produktion der entsprechenden **Genprodukte** nutzen konnte: Hier eröffnete sich zum ersten Mal die Möglichkeit, Proteine von therapeutischer Bedeutung in biotechnischen Verfahren zu erzeugen, die sich andernfalls nur äußerst mühsam oder in nicht ausreichender Menge isolieren ließen, oder überhaupt nicht zugänglich waren. Ein Novum war, dass Forscher, zuerst in den USA, eigene Unternehmen gründeten, um die Entwicklung gentechnischer Produkte voranzutreiben. Zu den ersten wirksamen Produkten, die gentechnisch hergestellt werden konnten, gehören Humaninsulin, menschliches Wachstumshormon sowie menschliche Interferone, die zur Bekämpfung viraler Infekte geeignet sind. Ehe wir zu Details übergehen, sollten wir die heutigen Kategorien und Ziele der Biotechnologie definieren [vgl. auch ref. [51].

Tabelle 3.1: Kategorien der Biotechnologie

Kategorie	Ziele
Rote Biotechnologie	Betrifft gentechnische Anwendungen auf medizinischem Gebiet und in der Gesundheitsvorsorge: Gewebe- und Organgewinnung; Entwicklung und Verabreichung von Pharmazeutika; molekulare Diagnostik und Genterapie.
Grüne Biotechnologie	Betrifft Pflanzen und Agrikultur: Pflanzliche Gewebekulturen; genetische Veränderungen von Pflanzen; Pflanzenzüchtung und Pflanzen-Hybride mittels ‚Zuchtmarkern‘; Entwicklung von Bioreaktoren und Biopestiziden.
Weißße Biotechnologie	Betrifft gentechnische Anwendungen in allen industriellen Prozessen: Entwurf und Entwicklung von umweltverträglichen Prozessen und Produkten bei kostensparenden Bedingungen; Entwicklung von Bioreaktoren für Zellkulturen oder zur Anzucht von Mikroorganismen; Herstellung von wertvollen Produkten durch transformierte Mikroorganismen, u.a. von homologen oder heterologen Metaboliten, Feinchemikalien, Antibiotika und anderen Arzneimitteln, Grundstoffen für Plastik- oder Textilproduktion, Produktion von nachhaltigen Energieträgern (Biotreibstoffe).
Blaue Biotechnologie	Betrifft Meeres- und Süßwasser-Organismen: Heranziehung solcher Organismen und ihrer Derivate, um Menge und Sicherheit von Nahrung aus dem Meer zu steigern; Kontrolle der unerwünschten Vermehrung schädlicher Wasserbewohner und Entwicklung neuer Medikamente oder Abwehrstoffe.
Graue Biotechnologie	Betrifft gentechnische Anwendungen, die direkt die Umwelt betreffen: Erhaltung der Biodiversität und Beseitigung von Abfallprodukten.

### 3.1.2 Anwendungen der Biotechnologie in der Medizin

Medizinische Anwendungen der Biotechnologie standen von Anfang an im Vordergrund (Abbildung 20).

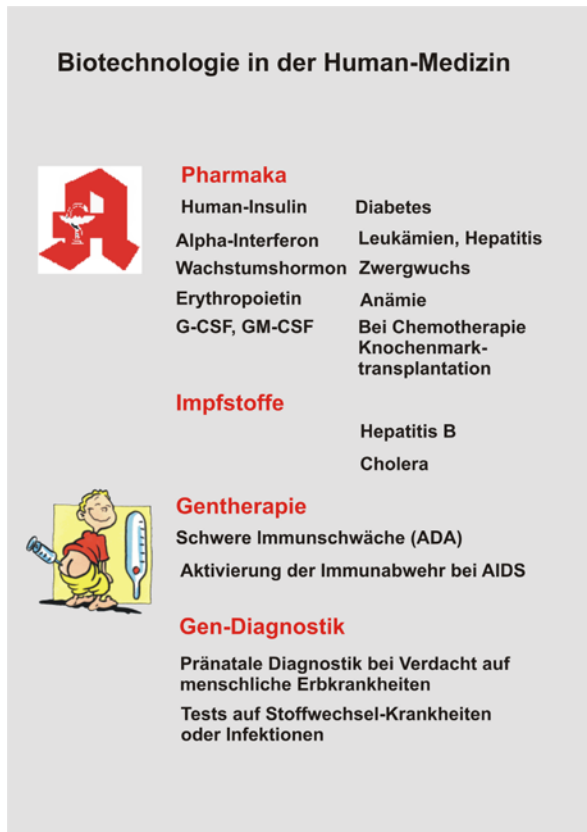


Abbildung 20: Anfängliche Anwendungen der ‚Weißen‘ und ‚Roten‘ Biotechnik.

Inzwischen werden fast alle Diabeteskranken mit biotechnisch hergestelltem Insulin behandelt, das weitaus verträglicher und sicherer ist als das zuvor verwendete tierische Produkt (modifiziertes Schweineinsulin). Insulin nimmt damit auch unter wirtschaftlichen Gesichtspunkten eine Spitzenstellung unter den heute verfügbaren gentechnisch hergestellten Medikamenten ein. Weitere gängige Therapeutika sind beispielsweise verschiedene Interleukine und andere Wachstumsfaktoren, die das Immunsystem stimulieren und bei Krebserkrankungen eingesetzt werden, oder Erythropoietin, ein renaler Wachstumsfaktor, der die Blutbildung stimuliert und bei chronischer Niereninsuffizienz lebenserhaltend wirkt. In die Palette gehören auch

verschiedene Impfstoffe, wie beispielsweise ein Hepatitis B-Vaccin, und zahlreiche Diagnostika zur Früherkennung von Virus- und Krebserkrankungen. Als Beispiel aus jüngerer Zeit ist die Gewinnung eines Protease-Inhibitors zur Bekämpfung des menschlichen Immunschwäche-Virus (HIV) zu nennen.

Biopharmazeutika können heute in Bioreaktoren aus rekombinanten Mikroorganismen (Bakterien wie *E. coli* oder Hefen) gewonnen werden, aus tierischen Zell-Linien oder sogar aus Pflanzenzellen. Wichtig sind dabei die folgenden Parameter in der Produktion: niedrige Kosten für Nährstoffe der Kulturen und geringer Energieverbrauch, möglichst geringes Volumen der Bioreaktoren, hohe Reinheit des Produktes (frei von mikrobiellen Verunreinigungen) [52].

Waren 1996 ca. 25% der weltweit neu auf den Markt genommenen Arzneimittel gentechnischen Ursprungs, sollten es im Jahre 2000 bereits über 60% sein: Die Palette umfasste Antidiabetika, Impfstoffe, Immun Stimulantien, Medikamente gegen Blutkrankheiten, verschiedene Hormone und viele andere, in geringerem Maße vertretene Produkte; der weltweite Umsatz belief sich 2000 auf etwa 2,2 Mrd. US\$, davon in Deutschland ca. 35% (900 Mio US\$). Mittlerweile sind viele neue Präparate hinzugekommen, vor allem im Zusammenhang mit den Ergebnissen aus dem Humangenom-Projekt.

**Tabelle 3.2** listet etwa 200 rekombinante Produkte auf, die bereits 2010 auf dem Markt waren; es handelt sich um in den USA und Europa biotechnologisch hergestellte Präparate, zurzeit dürfte es mindestens 100 Neuzulassungen geben. Nicht eingeschlossen sind Hersteller aus anderen Teilen der Welt. Jedoch ist anzumerken, dass sich die Zahl der Zulassungen in den USA und Europa durchaus unterscheiden, weil die jeweiligen Aufsichtsbehörden unterschiedlich bewerten. Für den Vertrieb der bis zum Jahre 2015 verfügbaren Biotherapeutika wird allein in den USA ein Umsatz von 320 Mrd. US\$ erwartet.

Unbestreitbar hat der therapeutische Wert gentechnischer Produkte die Akzeptanz dieser Seite der Gentechnik („Rote Biotechnologie“) ungeheuer gestärkt. Kaum ein Patient dürfte heute noch hinterfragen, auf welche Weise ein wirksames Medikament entstanden ist, solange es ihm hilft und frei von unerwünschten Nebenwirkungen bleibt. Wir wissen inzwischen, dass manche synthetischen Arzneimittel wegen ihrer Nebenwirkungen vom Markt genommen werden mussten; bei Biotherapeutika gab es solche Rücknahmen auch, aber nur in ganz wenigen Fällen.

Leser, die sich für Einzelheiten interessieren, sei eine umfassende Tabelle (englisch) unter ref. [53] empfohlen.

Erwähnen sollte man an dieser Stelle, dass skrupellose Anwender von (biotechnischen) Pharmaprodukten uns die Skandale des Doping beschert haben. Der Aufwand der Dopingkontrollen bei allen Arten von Wettkämpfen (national oder international), seien es Olympische Spiele, Tour de France, Weltmeisterschaften, o. dgl., ist technisch und finanziell enorm. Auf diese Weise wird Volksvermögen verschleudert, das andernfalls in der Gesundheitsvorsorge wesentlich effizienter eingesetzt werden könnte. Ich selbst habe die Dopingkontrollen zu den Olympischen Spielen in München 1972 mitverfolgen können, weil ein Team der Universität zu Köln unter Leitung meines Studienkollegen *Professor Donicke* mit den Untersuchungen in unserem Institut betraut war; er war selber erfolgreicher Radrennfahrer. Ich habe auch die Kosten für Geräte und Aufwand kennen gelernt. Heute sind die Analysemethoden immer weiter verfeinert, dadurch aber auch wesentlich kostspieliger geworden.

Tabelle 3.2: In den USA und Europa zugelassene Biopharmaka im Jahre 2010

Gruppe	Therapeutische Indikationen	Wirksame Produkte	Zahl der Präparate	Wirt*)
<b>Blutfaktoren</b>	Hämophilie A	Faktor VIII Antithrombin	8 5	
<b>Thrombolyse-Hemmer und Anticoagulantien</b>	Myocard-Infarkt Blutgerinnungsfehler/Venenthrombose Schwere Sepsis	Gewebe Plasmin Aktivator Anticoagulans Protein C	6 3 1	
<b>Hormone</b>	Diabetes Wachstumsanomalien (auch Turner-Syndrom) Fruchtbarkeitsanomalien Diabetes Typ 2 Osteoporosis Unfruchtbarkeit Schilddrüsenkrebs Hypoglykämie	Insulin Wachstumshormon (Somatotropin)  Follikel-stimulierendes Hormon Glukagon Parathormon, Calcitonin Lutropin Thyrotropin-±	16 13 1 5 3 4 1 1	H   H
<b>Wachstumsfaktoren</b>	Anämien Neutropenie  Wachstumsanomalien bei Kindern Mucositis bei hämatologischem Krebs Knochen-Diabetes, Neuropathien	Erythropoietin (EPO) G-CSF (Granulocyten Kolonie-stimulierender Faktor) IGF-1 (Insulin-ähnlicher Wachstumsfaktor) Keratinocyten- Wachstumsfaktor PDGF (Platelet-derived growth factor)	13 10 2 1 2	 H  H
<b>Interferone and Interleukine</b>	Chronische Hepatitis C (auch B); Krebs Multiple Sklerose Chronische Granulomatose Rheumatoide Arthritis Thrombozyten-Insuffizienz Nierenzellkrebs	Interferon-± Interferon-β Interferon-γ IL-1 (Interleukin 1) IL-11 (Interleukin 11) IL-2 (Interleukin 2)	10 4 1 1 1 1	
<b>Impfstoffe (Vaccine)</b>	Hepatitis A Hepatitis B (auch A) Kombi-Impfung (Diphtherie/Keuchhusten/Hep. B/Polio/Grippe/Tetanus) Prostata Krebs Gebärmutterhalskrebs Cholera Lyme (Borreliose) Meningokokken	Hepatitis A Antigen Hepatitis B Antigen Entsprechende Antigene  PAP-GM-CSF Papillomvirus Typ 16/18 (HPV) Vibrio cholerae ToxinB B.burgdorferi Oberflächenantigen Meningokokken Proetinkonjugat	1 5 6  1 1 1 1 1	H H H  H  H
<b>Monoklonale Antikörper</b>	Osteoporose (Postmenopause) Rheumatoide Arthritis Chronische lymphatische Leukämie Ascites Tumoren Psoriasis Macula-Degeneration (feuchte AMD) Hämoglobinurie Dickdarmkarzinom Multiple Sklerose Asthma Lymphome Akute lymphatische Leukämie Brustkrebs Prophylaxe gegen Abstoßung von Nierentransplantaten Diagnose von Lungenkrebs Diagnose von Melanomen Diagnose von Prostata-Adeno-Karzinom Diagnose von Dickdarmkrebs Diagnose von Infarkten Diagnose von Eierstockkrebs/Dickdarmkrebs Verhütung von Blutklumpung	Antikörper gegen RANKL Antikörper gegen IL-6; TNF-α Antikörper gegen CD 20  Antikörper gegen IL-12 + IL-23/LFA-1 Antikörper gegen VEGF-A Antikörper gegen C5 Komplement Antikörper gegen hEGFR Antikörper gegen Leukozatenintegrine Antikörper gegen IgE Antikörper gegen CD20 antigen Antikörper gegen CD33 Antikörper gegen HER2 Antikörper gegen IL-2  Antikörper gegen HMV-MAA Antikörper gegen PSMA Antikörper gegen CEA Antikörper gegen Herzmuskel-Myosin Antikörper gegen CA 125 Antikörper gegen PSR GPIIb/IIIa	1 4 2 1 2 1 1 2 2 1 4 1 1 1 1 2 2  1 1 1 1 1 2 2	
<b>Knochenbildungs-Proteine</b>	Posterolaterale lumbal-spinale Fusion Offene Schienbeinfrakturen Rückgrat-Verletzungen	BMP-7 (bone morphogenetic protein) BMP-2 BMP-2	1 4 1	
<b>Enzyme</b>	Pompe Krankheit (Glycogen-Speicher-Krankheit) Gaucher Krankheit Hunters Syndrom (Mucopolysaccharidose II) Adjuvans Fabry Krankheit (Galaktosidase-Defizienz) Hyperuricämie Cystische Fibrose Macula-Degeneration Cytomegalo-Virus Retinitis bei AIDS	±-Glucosidase Glucocerebrosidase Iduronate-2-sulfatase Hyaluronidase ±-galactosidase Uratoxidase DNase und Nukleinsäure-Derivate Spezifisch bindende Oligonucleotide für Vasculären Endothelwuchsfaktor Antisense Oligonucleotide	1 3 3 1 2 3 1 1 1	      H
<b>Weitere Proteine</b>	Cryopyrin-gekoppelte periodische Syndrome (CAPS) Thrombozyten-Defizienz Rheumatoide Arthritis	IL-1 Rezeptor  Thrombopoietin-Rezeptor Zytotoxisches T-ly.phozyten antigen 4	1  1 2	



	Congestive Herzinsuffizienz	Natriuretisches Peptid	1	
	Psoriasis	LFA-3 Protein	1	
	T-zell Lymphom der Haut		1	
	Blutgerinnsel	Thrombin-inhibitor	1	H

### 3.1.3 Transgene Organismen

Als transgene Organismen bezeichnet man solche, die durch Einbringen ‚fremden‘ genetischen Materials erzeugt werden (Abbildung 21); man bezeichnet sie auch als ‚genetisch veränderte Organismen‘ (GVO) oder im Englischen als ‚genetically modified organisms‘ (GMO) [54].

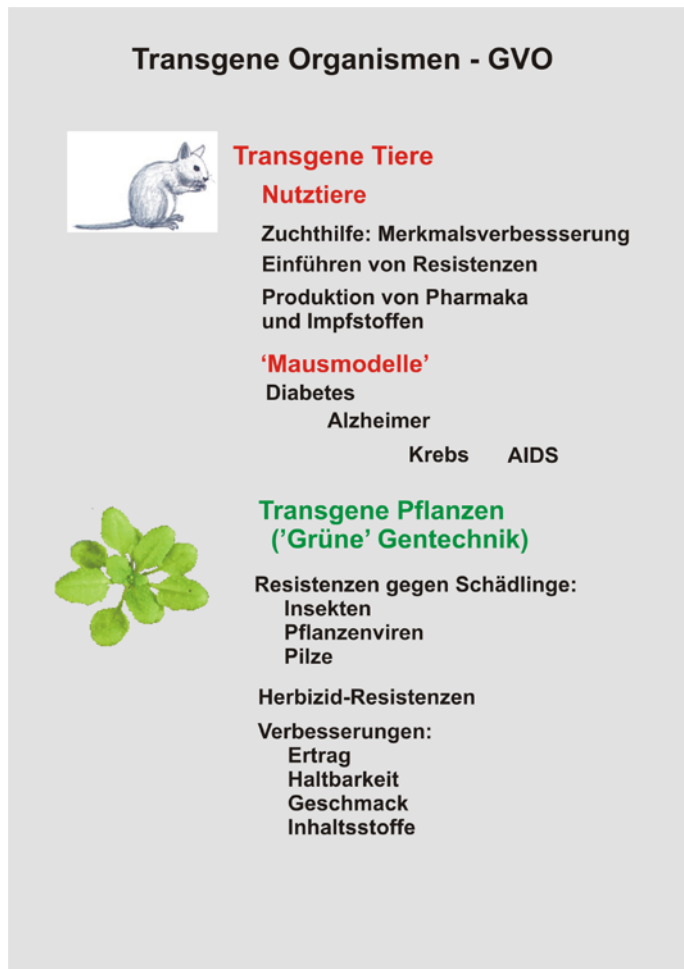


Abbildung 21: Nutztier und Pflanzen als Empfänger von fremdem Genmaterial.

#### 3.1.3.1 Transgene Tiere

Heute sind viele technische Detailfragen sowohl zum Gentransfer als auch zur gezielten Entfernung von Genen, nicht nur bei Mikroorganismen, sondern auch bei vielen höheren Organismen, bereits zufriedenstellend gelöst. Im Einklang mit den jeweilig geltenden Richtlinien, welche auch ethische Bewertungen von Projekten mit gentechnisch veränderten Organismen (GVO) erfordern, sind von biotechnisch einsetzbaren Mikroorganismen über so genannte Mausmodelle (z.B. ‚knockout‘ Mäuse, die bei der Erforschung von Genfunktionen helfen) bis hin zu transgenen Nutztieren [54, 55] sinnvolle Varianten durchgespielt worden.

Genetisch modifizierte Säugetiere stellen eine wichtige Kategorie unter den GVOs dar. *Ralph Brinster* und *Richard Palmiter* entwickelten schon in den 1980 Jahren Techniken zur Erzeugung transgener Mäuse, Ratten, Kaninchen, Schafe und Schweine und etablierten die ersten ‚Modelle‘ für genetische Krankheiten des Menschen, u.a. auch für die Entstehung von Karzinomen.

Heutige Entwicklungen beschäftigen sich mit folgenden sechs Themenkomplexen:

1. Erforschung menschlicher Krankheiten
2. Erzeugung von industriellen und/oder verbraucherfreundlichen Produkten
3. Erzeugung von therapeutisch nutzbarer Produkte, wie Biopharmaka oder Gewebe für Transplantationen
4. Verbesserung der Haustierhaltung (hypo-allergene Haustiere)
5. Verbesserung der Produktion von qualitätsreichen Nahrungsmitteln (bei Fischen, Schweinen, Kühen)
6. Verbesserungen für die Gesundheit von Tieren (Krankheitsresistenzen)

Transgene Tiere werden in der biomedizinischen Forschung als vielversprechende Modelle eingesetzt. Hervorzuheben sind hier vor allem ‚Mausmodelle‘ (‚knockout‘-Mäuse), deren Zahl heute weit in die Tausende geht; mehrere Zentren sind weltweit für ihre Haltung und Erforschung eingerichtet worden. 2007 wurde der Nobelpreis für Medizin oder Physiologie an *Mario R. Capecchi*, *Sir Martin J. Evans* und *Oliver Smithies* für ihre "Entdeckung von Prinzipien, spezifische Genmutationen in Mäuse mithilfe embryoner Stammzellen einzuführen" vergeben. Gleichzeitig beinhalteten diese Techniken auch die Möglichkeit, spezifische Mausgene aus dem Genom gezielt zu entfernen oder durch mutierte Gene zu ersetzen. Sie haben wesentlich dazu beigetragen, genetische Krankheiten beim Menschen zu erforschen, geeignete Medikamente oder Gentherapien zu entwickeln, oder Transplantate anzupassen.

Auch andere Säugetiere (wie Schweine, Schafe und Ratten) wurden gentechnisch – auf ähnliche Art wie Mäuse – verändert und gezüchtet. Transgene Schweine dienen dazu, Organtransplantate zu studieren, zum Beispiel zur Regeneration lichtempfindlicher Zellen im menschlichen Auge oder neuronaler Gehirnzellen; weiterhin zur Etablierung von Stammzellen in der regenerativen Medizin oder zur Züchtung von Gewebetransplantaten. Japanische Forscher haben 2009 sogar transgene Seidenäffchen hervorgebracht, um Parkinson, Huntington und amyotrophe laterale Sklerose an einem Primaten zu erforschen.

Weiterhin werden transgene Tiere zur Produktion von Biotherapeutika oder menschlichen Proteinen (‚gene pharming‘) eingesetzt. 2009 ließ die *U.S. Food and Drug Administration* das erste, aus Ziegenmilch gewonnene Medikament zu, ein Anticoagulans, welches Blutgerinnsel verhindert. Seit 2001 gibt es chinesische Kühe, die Milch produzieren, die die gleiche Zusammensetzung wie menschliche Muttermilch besitzt; ähnliche Entwicklungen geschahen in Argentinien und Neuseeland.



Abbildung 22: Der Aquarienziebling „Glofish“

[http://www.glofish.com/images/glofish\\_005.jpg](http://www.glofish.com/images/glofish_005.jpg)

Auch transgene Fische wurden erzeugt, zum Beispiel fluoreszierende Sorten von *Danio rerio* (Abbildung 22), einem Aquarienziebling. Bisher nicht geklärt ist, ob transgene Lachse, gezüchtet seit 2010 in Aquakulturen, die doppelt so groß werden wie Wildlachse, zum Verzehr freigegeben werden [56].

### 3.1.3.2 Gentherapie beim Menschen

Wir müssen vorausschicken, dass gegenwärtige Gentherapien auf somatische (nicht-reproduktive) Zellen beschränkt bleiben, so dass transferierte Gene nicht auf die folgende Generation übertragen werden können. Sogenannte „Keimbahn-Gentherapie“ bleibt höchst umstritten und wird für die nahe Zukunft keine Chance haben.

Gentherapie wurde bereits kurz nach Etablierung gentechnischer Verfahren versucht. Der erste Patient (1968) war ein kleines Mädchen mit einer bestimmten Form einer genetischen Krankheit, genannt ‚Schwere Kombinierte Immunschwäche‘ (severe combined immunodeficiency - SCID). Ursache war der Ausfall des Adenosindeaminase-Gens (ADA), welches mit Hilfe einer Knochenmarks-Transplantation von manipulierten hämatopoetischen Stammzellen wiederhergestellt wurde und vorübergehend funktionierte; das Kind überlebte aber nur kurze Zeit und musste in steriler Atmosphäre aushalten. Im Laufe der Zeit wurden bessere Bedingungen für die biotechnische Manipulation der Stammzellen und der Transplantation ausgearbeitet, vor allem, um Problemen im Bereich von Infektionen zu begegnen. 2007 war die Überlebenschance von transplantierten SCID-Patienten auf 70-80% gestiegen; die Fortschritte beschreibt ein Artikel (englisch) aus diesem Jahre [57].

Erfolge bei der somatischen Gentherapie stellen sich immer noch zögerlich ein. Zwar sieht das Prinzip, nämlich die Möglichkeit Krankheiten, die ihre Ursache in fehlenden oder aberranten Genen haben, durch das Einschleusen von ‚gesunden‘ Genen auszugleichen oder sogar die schadhafte Gene zu ersetzen, überzeugend aus, jedoch sind die technischen Schwierigkeiten und auftretende Komplikationen vielfältig: Die eingebrachten Gene müssen in richtiger Dosierung und an den ‚richtigen‘ Ort im Genom (das heißt, auch in den originären Zelltyp) integriert werden, um wirksam zu sein; es müssen Vehikel für den Gentransfer gefunden werden, deren DNA sich später nicht, wie es von viraler DNA bekannt ist, unliebsam verselbständigen kann. An tierischen ‚Modellen‘, vor allem dem der Maus (s. weiter unten), haben sich gentherapeutische Erfolge aufzeigen lassen; in vielen

Fällen jedoch haben sich immer wieder beträchtliche Nebenwirkungen herausgestellt. Obwohl es 1995 weltweit schon über 100 genehmigte Gentherapien gab, wird man im Einzelfall immer noch die Chancen gegen Risiken abwägen müssen. Die inzwischen gesammelten Erfahrungen zeigen jedenfalls, dass die Entwicklung wirksamer Gentherapien Zeit und beträchtlichen Aufwand erfordert, und dass der frühe Enthusiasmus mancher Forscher zu weit gehende Erwartungen geweckt hat.

Neuere Ansätze befassen sich mit der Gentherapie folgender genetischer Krankheiten: Cystische Fibrose (Mukoviszidose), Sichelzellanämie, Parkinson'sche Krankheit, Muskeldystrophie, Diabetes und Leberkrebs.

Die Mukoviszidose (auch: Cystische Fibrose) ist bei Vertretern der kaukasischen Rasse eine häufige genetische Stoffwechselstörung, die autosomal-rezessiv vererbt wird. Unter den Nachkommen zweier ‚Träger‘ (die selber gesund scheinen) wird immer nur einer von vier Nicht-Träger, zwei von vier werden wieder Träger, und einer von vier ist von vornherein krank. Die Funktionsstörung betrifft ein Gen (cystic fibrosis transmembrane conductance regulator – CFTR), das den Transport von Chlorid im Körper regelt und in die Produktion von Sekreten, Schweiß oder Schleim eingreift. Ist dieses Gen mutiert, kommt es je nach Art der Mutation von äußerlich therapierbaren bis zu gravierenden Störungen, meist Unfähigkeit der Schleimentsorgung, betroffen sind Lungen und Atemwege. Hinzu kommen Probleme der inneren Sekretion (Pankreas), Infertilität bei 97% der betroffenen Männer, höhere Anfälligkeit für Infektionen über die Atemwege [58]. Mukoviszidose kann über pränatale Diagnostik vorhergesagt werden. Sie konnte in früheren Zeiten unter Umständen schon im Kindesalter zum Tod führen, denn sie ist bisher nicht heilbar. Wohl aber haben sich die Behandlungsmethoden in den vergangenen 70 Jahren zunehmend verbessert. Es gibt biotechnisch hergestellte Medikamente, die sowohl die ‚Restwirkung‘ mutierter Gene erhöhen als auch solche, die zähen Schleim verflüssigen helfen. Unbedingt notwendig ist die Senkung von Infektionsrisiken über die Atemwege durch (prophylaktische) Gabe von geeigneten Antibiotika. Ein ‚Genersatz‘ bei Probanden über eingebrachte ‚gesunde‘ Gene wird zurzeit untersucht.

Forschungen, die anderen, oben erwähnten, unheilbaren genetischen Krankheiten durch Gentransfer zu kurieren, sind in vollem Gange. Bestrebungen, die humane Gentherapie weiter zu entwickeln, haben bereits nach Europa übergegriffen [59].

### **3.1.3.3 Transgene Pflanzen**

Auf lange Sicht wird aus verschiedenen Gründen die Bedeutung gentechnisch veränderter Pflanzen zunehmen [60]. Eines der ersten Argumente, der Anbau von GVO Pflanzen für die Ernährung sei wegen der wachsenden Weltbevölkerung unbedingt notwendig, ist nach den bisherigen Erfahrungen weitgehend hinfällig, denn Entwicklungsländer können sich den Anbau solcher Pflanzen bis heute selten leisten; in vielen von ihnen bestehen auch noch keine gesetzlichen Regelungen. Hingegen haben gerade Entwicklungsländer den Anbau von GVO Nutzpflanzen – vornehmlich Baumwolle – weit vorangetrieben.

Führende Agrarfirmer haben längst viele ‚transgene‘ Nutzpflanzen (z.B. Tabak, Mais, Tomate, Kartoffel, Sojabohne, Raps, Melone, Baumwolle, Reis, Gerste) erzeugt und stellen weltweit Saatgut bereit, unter anderem Varianten, die durch gezieltes Einschleusen von geeigneten Genen gegen spezifische Schädlinge (Insekten, Pilze, Bakterien, pflanzliche Viren) und/oder eingesetzte Herbizide resistent gemacht wurden, oder gentechnisch veränderte Varianten, die es ermöglichen, die Haltbarkeit oder den Ertrag der pflanzlichen Produkte zu steigern (Abbildung 23).

Transgene Nutzpflanzen			
Apfel	Himbeere	Pfeffer	Spargel
Aubergine	Kartoffel	Pflaume	Süßkartoffel
Baumwolle	Kiwi	Preiselbeere	Tabak
Blumenkohl	Kohlarten	Raps	Tomate
Erbsen	Luzerne	Reis	Walnuss
Erdbeere	Mais	Roggen	Weintraube
Fichte	Meerrettich	Salat	Weizen
Flachs	Papaya	Sellerie	Zuckerrohr
Gurke	Pappel	Soja	Zuckerrübe

<b>Gentransfer</b>
(1) Über modifizierte DNA-Vehikel aus Bodenbakterien, die sonst Gene in ihre Wirtspflanze einschleusen und sie zur Bildung von Wurzelhalsgallen (Tumore) veranlassen; auch über "Knöllchenbakterien" von Leguminosen
(2) "Beschuß" von Pflanzen mit winzigen Goldpartikeln, die mit DNA beschichtet sind
<b>Beispiele</b>
<b>Verbesserung der Beschaffenheit von Feldfrüchten</b>
Kartoffel: 60 % mehr Stärke; größere Knollen
Mais: günstigere Aminosäurezusammensetzung
<b>Verbesserung der Haltbarkeit von Früchten</b>
Tomate: Verzögern der Reifung
Erdbeere: Schutz vor Frost
<b>Schutz gegen Insektenbefall</b>
Gene von <i>Bacillus thuringiensis</i> (Bt) Toxinen (ungiftig für Säugetiere)
Baumwolle: Raupenfraß ('bollworm')
Kartoffel: Colorado (Kartoffel-) Käfer
Mais: Mais-Schädlinge
<b>Resistenzen gegen Pilzbefall</b>
Mais:
<b>Resistenzen gegen Virusbefall</b>
Kartoffel: Y-Virus
Tabak: Tabak-Mosaikvirus
Tomate: Bushy Stunt Virus
Weinrebe: 'Reisig-Krankheit'
<b>Resistenzen gegen Breitband-Herbizide</b>
Tomate, Sojabohne, Baumwolle, Raps: Glyphosat (Hemmstoff für pflanzenspezifische Enzyme; unbedenklich für Tiere)

Abbildung 23: ‚Grüne Gentechnik.‘

Dabei muss natürlich die besondere Problematik mit gentechnisch veränderten Organismen (GVO) auf dem Pflanzensektor berücksichtigt werden: Während Mikroorganismen in der Produktion ‚eingesperrt‘ sind und Rückstände effizient beseitigt werden können, müssen Pflanzen freigesetzt werden. Daher muss durch besondere Sicherheitsvorkehrungen gewährleistet sein, dass schädliche Auswirkungen sowohl auf die Umwelt als auch auf Mensch und Tier als ‚Verbraucher‘ ausgeschaltet bleiben. Eine berechtigte Befürchtung ist auch, dass sich die eingebrachten Gene in unerwünschter Weise auf andere Pflanzen übertragen. So wird in den USA die Zulassung und Produktion transgener Pflanzen seit 1994 unter strenger Aufsicht der *Food and Drug Administration* (FDA) betrieben, um



eventuelle Risiken zu minimieren. Dennoch wurden auch hier vor einiger Zeit kritische Stimmen der Verbraucher laut, wie sie – in Europa und besonders in Deutschland – immer noch gegenüber der ‚grünen Gentechnik‘ vorherrschen. Regelungen in anderen Ländern und in der EU werden in ref. [60] vorgestellt; dort werden auch ethische Aspekte angesprochen [61].

Im Jahre 2000 waren weltweit etwa 90 gentechnisch veränderte Pflanzen zugelassen und 44 Mio Hektar wurden damit bebaut, davon in den USA 68%, in Argentinien 23%, und in Kanada 7%. Der Gesamtertrag aus der entsprechenden Produktion von Soja (36%), Baumwolle (16%), Raps (11%) und Mais (7%) lag dabei über 3 Mrd. US\$. In Europa dagegen (in Deutschland, Frankreich, Spanien) wurde der Anbau gentechnisch veränderter Pflanzen (u.a. Mais, Raps, Zuckerrübe, Kartoffel, Tomate) erst seit 1998 und nur zu Versuchszwecken erlaubt und machte unter 20.000 Hektar aus.

Bereits 2010 war der Anbau von GV Pflanzen weltweit auf 148 Mio Hektar nutzbaren Ackerlandes (11% allen verfügbaren Ackerlandes) gestiegen; 2012 sollte er auf 170 Mio Hektar ansteigen. Allein für Saatgut lag die Marktgröße transgener Pflanzen bei 6,9 Mrd. US\$, für 2025 wird sie auf 50 Mrd. US\$ geschätzt. Interessierte Anwender in der EU und in Deutschland warten jetzt auf die Novellierung EU-weiter Richtlinien, die sich wegen der geringen Akzeptanz gentechnisch veränderter Lebensmittel durch die europäischen Verbraucher allerdings schwierig gestaltet [62, 63].

### 3.1.4 Weiße Biotechnik

Die **Weiße Biotechnologie**, besser vielleicht Industrielle Biotechnologie genannt, setzt biotechnologische Methoden für industrielle Produktionsverfahren ein. Dazu werden biologische und biochemische Prozesse in technische Anwendungen übertragen. Als produzierende Organismen werden dafür gentechnisch modifizierte Mikroorganismen – hauptsächlich Bakterien oder Hefen – herangezogen. In den vergangenen Jahren haben sich Zahl und Einsatzmöglichkeiten dieser Organismen ständig erweitert, ebenso wurde die Entwicklung biotechnologischer Methoden und Anwendungen vorangetrieben. Von den anfänglich auf die Erzeugung medizinisch wertvoller Produkte beschränkten Unternehmungen, die bereits in Abschnitt 3.1.2 berührt wurden, hat sich die Industrie auch auf großtechnisch hergestellte Produkte ausgedehnt [64, 65, 66]:

- (1) Optimierung von Produktionsverfahren, zum Beispiel für Grund- und Feinchemikalien;
  - (2) Reduzieren der Rohstoffabhängigkeit, zum Beispiel durch Nutzung nachwachsender Rohstoffe statt fossiler Rohstoffe;
  - (3) Reduzierung der Energie- und Entsorgungskosten, zum Beispiel Ersetzen chemischer Verfahren durch biologische;
  - (4) Entwicklung neuer Produkte und Systemlösungen mit hohem Wertschöpfungspotenzial, zum Beispiel durch Nutzbarmachung von biologischen Stoffwechselwegen mit gentechnischen Methoden.
- Im Folgenden werden beispielhaft einige Anwendungen besprochen; wir wählen die Einteilung aber nicht nach Produktkategorien, sondern nach den hauptsächlich eingesetzten Organismen.

#### 3.1.4.1 Hefen

Vor allem *S. cerevisiae*, die Brauerei- oder Bäckerhefe, wird heute als geeignete ‚Zellfabrik‘ eingesetzt. Schließlich ist sie für die Menschheit seit Jahrtausenden als hilfreicher Organismus beim Backen, Brauen oder in der Weinbereitung unabdingbar und für den Verzehr selbst völlig unschädlich. Kein Wunder also, dass Hefe oder mit ihrer Hilfe hergestellte Produkte als wertvolle **Nahrungsergänzungsmittel** geschätzt werden (u.a. auch als Hefe-Extrakt erhältlich) oder als Futterhefe in großen Mengen an Tiere verfüttert werden kann.

Durch biotechnisch gelenkte Veränderungen lässt sich Hefe so manipulieren, dass sie wertvolle Stoffwechselprodukte in einem abschöpfbaren Ausmaß liefert. Speziell getrimmte Stämme können folgende Stoffe ausschütten: *Isoprenoide und Carotenoide* (z.B.  $\beta$ -Carotin als Vorläufer des Vitamin A; *Astaxanthin*, das als Futtermittelzusatzstoff (E 161j) zum Fischfutter bei der Erzeugung von Speisefischen verwendet wird, um den ansonsten weißfleischigen Regenbogenforellen ein lachsrotes Fleisch anzufüttern; sie werden dann als Lachsforelle vermarktet).

*Artemisinin* wird weltweit zur Behandlung von Infektionen mit multiresistenten Stämmen von *Plasmodium falciparum*, dem Erreger der Malaria, eingesetzt. *Taxol* (oder Paclitaxel genannt) wird als Arzneistoff zur Behandlung verschiedener Krebsarten (z. B. Brustkrebs) benutzt. Die Wirkungen von Biotin (Vitamin H), *Vitamin C*, oder *Cortisol* dürften allgemein bekannt sein. Im Rahmen des Screening-Programms des *National Cancer Institute* hat man *Resveratrol* in 72 Pflanzenarten (in roten Weintrauben, Himbeeren, Maulbeeren, Pflaumen, Erdnüssen) gefunden. Studien haben positive Effekte von Resveratrol bei Arteriosklerose, Herzkrankheiten, Alzheimer-Krankheit, Arthritis und manchen Autoimmunkrankheiten gezeigt.

Biotechnisch in großem Maßstab in Hefen erzeugte Produkte betreffen monomere Bausteine für die **Kunststoff- und Polymerherstellung**. Ziele sind dabei, erstens petrochemische Verfahren zur Herstellung bekannter Kunststoffe (die meist nicht biologisch abbaubar sind) durch biotechnologische Verfahren zu ersetzen oder neue Polymere mit neuen Eigenschaften zu entwickeln; zweitens biologisch abbaubare Polymere (Biokunststoffe) zu generieren. Zur ersteren Kategorie gehört Polymilchsäure (Polylactid, PLA): In Hefe lassen sich große Mengen von Milchsäure produzieren, die als Ausgangsmaterial gebraucht wird. Versuche, Glycerin (in Hefe) in 1,3-Propandiol umzuwandeln, sind im gang; dieses stellt einen bedeutenden Rohstoff in der Herstellung von Polymeren, Schmiermitteln oder Kosmetika dar.

Die zweifellos bedeutendsten biotechnisch in Hefe erzeugten Produkte sind Bioalkohole (vorwiegend **Bioethanol**), die durch Fermentation der Biomasse von Pflanzen oder Mikroalgen hergestellt werden. Große Kontingente an zuckerhaltiger Biomasse liefern Zuckerrohr oder Zuckermelasse, Zuckerrüben, Stärke, Weizen, und in der ‚zweiten Generation‘ auch Pflanzenabfälle, Kompost oder Abfälle von Nahrungsmitteln. Während Brasilien der Vorreiter in der Bioethanol-Produktion ist, wird Bioethanol in Deutschland (auch als Beimischung) kaum angenommen. Nicht nur bei uns, sondern weltweit werden

heftige Debatten um Biotreibstoffe generell geführt, denn auch die Auswirkungen der Herstellung von Biodiesel, Biobutanol und anderen aus Biomaterial erzeugten Treibstoffen bleiben problematisch. Der Leser mag sich selber Einblicke in diese Thematik verschaffen [67, 68].

### **3.1.4.2 Bakterien**

#### **3.1.4.2.1 *Escherichia coli* [69]**

Abgewandelte Zellen von *E.coli* werden zur Produktion von Impfstoffen und vielen anderen Biopharmazeutika eingesetzt, wie schon in Abschnitt 3.2.1 erwähnt. Sie spielen eine wichtige Rolle in der Produktion von Enzymen, die in der Lebensmittelindustrie Verwendung finden: Sie optimieren Fette und Proteine, helfen bei der Herstellung von Formfleisch, werden in der Eis- und Getränkeproduktion eingesetzt oder zur besseren Konservierung von Lebensmitteln benutzt.

Waschmittel sind heute weitgehend frei von Oxidantien. Stattdessen werden ihnen biotechnisch erzeugte Lipasen, Proteasen und Amylasen beigegeben, welche Verschmutzungen aller Art auflösen können.

Gentechnisch wurden *E.coli*-Stämme so modifiziert, dass sie für viele Aufgaben in der Umweltsanierung (z.B. in Kläranlagen) geeignet sind. Neuerdings wurden reichlich komplexe *E.coli*-Stämme geschaffen, die in der Lage sind Kohlenwasserstoffe zu erzeugen wie sie natürliche Treibstoffe enthalten („Benzin aus Bakterien“): Diesen Bakterien wurden Gene aus dem Kampferbaum, Bodenbakterien und blau-grünen Algen eingepflanzt. Der ‚letzte Schrei‘ sind *E.coli* Zellen, die licht-erzeugende Enzymsysteme (Luciferase) enthalten, die entweder direkt oder in Pflanzen eingebracht werden [70]. Diese sollen tagsüber Sonnenlicht einsammeln und die Energie als Lumineszenz im Dunkeln abgeben.

#### **3.1.4.2.2 *Corynebacterium***

Corynebakterien sind meist für den Menschen pathogen. Nicht-pathogene Spezies haben sich seit geraumer Zeit als bedeutende ‘Zellfabriken’ für die Produktion von Aminosäuren (Glutamat, Lysin, Threonin), Nukleotiden und anderen ernährungswichtigen Faktoren etabliert; Glutamat spielt eine Rolle in der Herstellung von Sojasauße oder Yoghurt. Sie werden eingesetzt zur Bioconversion von Steroiden, in der Käsebereitung und zur Produktion von Enzymen. Einige Arten können auch den Antibiotika ähnliche Verbindungen, wie Bakteriozine (Antitumor-Wirkstoffe), synthetisieren.

## 4 Genfunktionen, Gengeflechte und Netzwerke

### 4.1 Genomics - Die gesamten Gene und ihre Funktionen in einem Organismus

Die Bibliothek der Genome ist in letzter Zeit durch das ‚Buch der Humangene‘ und anderer Genomprojekte gewaltig gewachsen, doch gedruckt werden diese Millionen von Seiten wohl nie. Nur eigens auf die riesige Informationsmenge zugeschnittene Computerprogramme können beim Lesen und bei ihrer Interpretation helfen. Erstellung, Integration und Verfügbarmachen von Datenbanken, vor allem aber die computergestützte Analyse der Daten, stellt die Bioinformatik vor eine nicht geringe Aufgabe. Wie schon angedeutet, hat die ‚*in silico*‘-Analyse viel dazu beigetragen, Genfunktionen zu identifizieren, und sie wird es auch weiterhin tun. Spezielle Programme wurden entwickelt, um die automatische Annotation von Genprodukten voranzutreiben; denn erst mit der Erschließung einer Funktion lassen sich fundierte Aussagen über die Bedeutung z.B. eines Proteins machen. Diese Möglichkeit hat uns jedoch nicht der Notwendigkeit enthoben, nach experimentellen Ansätzen zu suchen, die Paralleluntersuchungen an vielen Genen gleichzeitig und den von ihnen exprimierten Proteinen in großem Umfange erlauben. Dabei wurde sofort deutlich, dass Massendurchsatz (ähnlich wie bei der Sequenzierung) nur mit Hilfe von Mikrotechniken (Nanotechniken im Massendurchsatz) erreicht werden kann.

Eine grundsätzliche Möglichkeit Genfunktionen experimentell zu erfassen, besteht darin, dass man gezielte Deletionen in das Genom einführt und damit das zu charakterisierende Gen aus dem Genom entfernt. Anschließend können dann die Auswirkungen hinsichtlich Wachstum, Erscheinungsbild oder Verlust bestimmter Stoffwechselwege untersucht und daraus auf die Genfunktion zurückgeschlossen werden. Im Sinne einer umfassenden Analyse sind beispielsweise komplette Sammlungen von Deletionsmutanten der Hefe (also etwa 6000) erzeugt und analysiert worden, in denen jeweils ein definierter Genort eliminiert wurde. Auch bei der Maus – gewissermaßen als tierischem Modellsystem - ist die gezielte Genzerstörung (*knock-out* Mäuse) in großem Umfang fortgeschritten, obwohl hier der Aufwand beträchtlich größer ist.

#### 4.1.1 Genomics mithilfe der Mikrochip-Technik

Ein Verfahren mit einer zunächst anderen Zielrichtung, die so genannte Mikrochip-Technik, wurde sehr früh für Hefe etabliert [71]. In der Tat gestattet diese Technik heute nicht nur, die Expression vieler Gene gleichzeitig zu erfassen, sondern diese in den verschiedenen Wachstumszuständen einer Zelle zu vergleichen, deren ‚Expressionsprofile‘ in zeitlicher Abfolge aufzuzeichnen, oder – vor allem im Hinblick auf höhere Organismen – unterschiedliche Zelltypen zu charakterisieren, in denen schließlich nur bestimmte Gene aus dem vorhandenen Genrepertoire gezielt an- oder abgeschaltet werden.

Die Mikrochip-Technik lässt sich am Beispiel der Hefe gut illustrieren. Zunächst werden Proben aller 6000 Gene, die man z.B. mit Hilfe der PCR gewinnen kann, in Form eines Rasters auf einem nur

0,324 cm<sup>2</sup> großen Träger fixiert. Lässt man dann einen solchen Mikrochip unter geeigneten Bedingungen mit der gesamten Boten-RNA, die man aus einer Zelle isoliert und durch Fluoreszenzfarben markiert hat, reagieren, so wird jedes der darin enthaltenen Moleküle aufgrund spezifischer Basenpaarung nur an die zugehörige Genprobe binden. Darüber hinaus lässt sich an der Stärke des Fluoreszenzsignals ablesen, welche Menge an RNA jedes einzelne Gen ursprünglich produziert hat. Da man allerdings keinen absoluten Standard zugrunde legen kann, beschränkt man sich auf einen Vergleich, indem man das zu messende Expressionsprofil jeweils auf einen internen Standard bezieht: Die RNA aus einer Hefekultur, die unter definierten Bedingungen in einem nährreichen Medium gewachsen ist, wird beispielsweise mit einem grünen Fluoreszenzfarbstoff markiert, diejenige aus der Versuchskultur dagegen mit einem roten. Führt man die oben beschriebene Reaktion nun mit einem Gemisch beider RNA-Präparationen aus, so konkurrieren die verschieden markierten Moleküle um den gleichen Partner, und es wird das vergleichsweise in höherer Menge vorliegende ‚gewinnen‘. Wie dann ein solches Muster von Fluoreszenzsignalen aussieht, erklärt das folgende Schema (Abbildung 24).

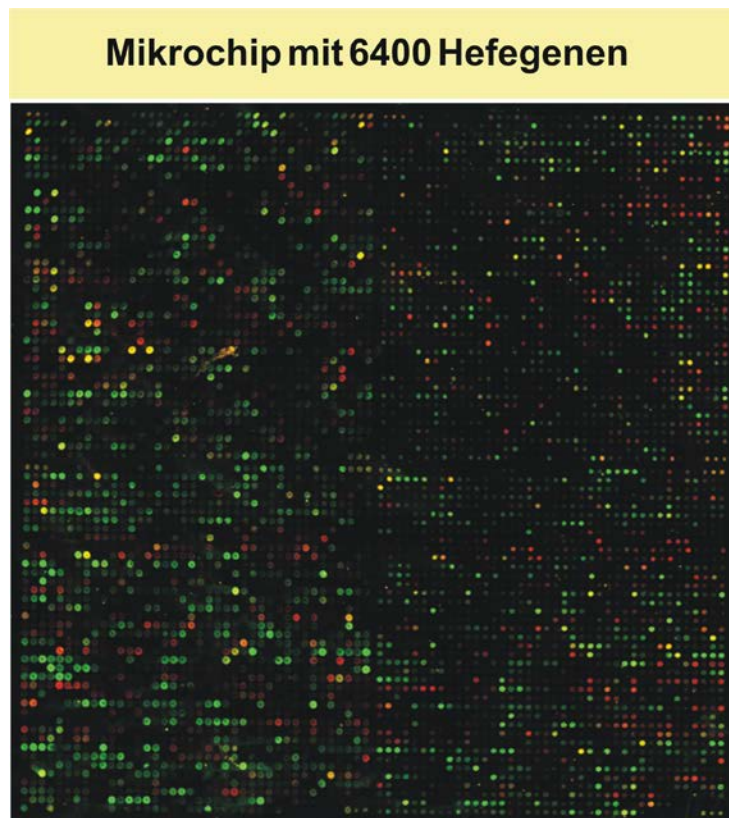


Abbildung 24: Ein Mikrochip zur Funktionsanalyse von Hefegenen [72].

DNA-Kopien aller 6400 Hefegene wurden mit Hilfe der Polymerase-Kettenreaktion (PCR) hergestellt und punktförmig mit Hilfe eines Roboters auf einer Glasmatrix von 18 x 18 mm Größe fixiert ('micro-arrays'). In diesem Falle wurde der Mikrochip benutzt, um Auskunft darüber zu erhalten, welche Hefegene an- bzw. abgeschaltet werden, wenn das Nährmedium an Glukose verarmt; die Hefe stellt dann nämlich ihre Energieversorgung von der durch aerobe (Sauerstoff benötigende)

Verwertung des Zuckers auf anaerobe (Sauerstoff unabhängige) Verwertung des mittlerweile produzierten Ethanols um. Dazu wurde zu Beginn des Experimentes aus einer Probe der angesetzten Kultur die gesamte Population der messenger-RNAs isoliert, in DNA zurückverwandelt und dabei alle diese Moleküle mit einem grünen Fluoreszenzfarbstoff markiert. In gleicher Weise wurde eine nach etwa neun Stunden gezogene Probe aus der wachsenden Kultur behandelt, nur wurde diesmal eine Markierung mit einem roten Fluoreszenzfarbstoff gewählt. Anschließend wurden gleiche Mengen der unterschiedlich markierten DNA-Sonden in einer geeigneten Lösung gemischt und der Mikrochip darin bei erhöhter Temperatur längere Zeit 'gebadet': Unter diesen Bedingungen 'hybridisieren' die farbigen



DNA-Moleküle nur mit dem Gen in der Matrix, aus dem sie einmal hervorgegangen sind: sie bleiben aufgrund der identischen Sequenzen gewissermaßen daran kleben. Nun kann man die in jedem Punkt der Matrix sichtbaren **relativen** Fluoreszenzen mit Hilfe eines automatischen konfokalen Scan-Mikroskops optisch erfassen. Rote 'Punkte' verschiedener Intensität im Muster kennzeichnen diejenigen Gene, die bei der Umstellung des Wachstums von Glukose auf Ethanol angeschaltet werden. Gene, die abgeschaltet werden, sind durch grüne Punkte verschiedener Intensität gekennzeichnet. Bei Genen, die unter beiden Bedingungen etwa gleich arbeiten, halten sich rote und grüne Fluoreszenz die Waage, so dass die Punkte wegen der Überlagerung gelb erscheinen. Da man zu jedem Punkt das zugehörige Gen angeben kann, ließ sich genau festlegen, von welchen der insgesamt 1740 betroffenen Gene während der Umstellung die Aktivität erhöht oder erniedrigt wurde, und um welchen Faktor.

Dieses Muster wird sich in dem Maße verändern, wie sich das **Expressionsprofil** in der Versuchskultur ändert. Eines der frühesten Experimente (1997) bestand darin zu zeigen, welche Gene in einer Hefekultur bei Verarmung des Mediums an Glukose an- oder abgeschaltet werden. Mit ähnlichen Ansätzen konnte man ermitteln, welche Hefegene zeitabhängig während des Wachstumszyklus oder des Sporulationsprozesses von Hefe in ihrer Expression verändert werden. Da der Zellzyklus in allen eukaryotischen Zellen ähnlich abläuft wie in Hefe oder Umstellungen des Stoffwechsels grundsätzlich mit Veränderungen von Genaktivitäten einhergehen, waren und sind die an Hefe gewonnen Erkenntnisse ein grundlegender Beitrag zum Verständnis der Zellregulation. Man erlaube mir, aus meinem Arbeitsgebiet einige Fortschritte anzuführen, die Gebrauch von der Mikrochip-Technik machen: „Messung von Wechselwirkungen zwischen DNA und Proteinen“ [73]; „Beschreibung aller Hefeproteine und Mikroarrays“ [74]; „Die genetische Landschaft einer Hefezelle“ [75].

Die Mikrochip-Technik hat sich in der Folgezeit zu einer Routine-Methode, vor allem auch im Verbund mit dem Humangenomprojekt, entwickelt. Sie wurde seit ihrer Einführung hier in vielen Studien verwendet, beispielsweise, um Genmutationen aufzufinden, um Expressionsprofile in verschiedenen Zelltypen zu vergleichen oder um Unterschiede der Genexpression von Krebszellen gegenüber gesunden Zellen aufzuzeigen. Solche Mikrochips werden heutzutage in vielen Varianten auf dem Biotechnologiemarkt angeboten.

#### 4.1.2 Andere Sparten von „-omics“

Wie das Genom die Gesamtheit der Gene eines Organismus beschreibt, so soll das ‚Proteom‘ den vollständigen Satz an Proteinen, der in einer Zelle in jeder Phase ihres Lebens exprimiert wird, erfassen [76, 77]. Ein systematischer Ansatz, wie bei den Genomprojekten, wird hier wegen der vom Zellzustand abhängigen Fluktuation der Proteinprofile wesentlich schwieriger; zudem waren die verfügbaren Methoden anfangs kaum auf einen Massendurchsatz zugeschnitten. Das ‚klassische‘ Verfahren, welches zelluläre Proteine auftrennen kann, ist die **zweidimensionale Gelelektrophorese**, die mehr oder weniger ‚punktförmige‘, charakteristische Muster der Proteine erzeugt. Anschließend

kann man solche Protein-Muster, z.B. aus verschiedenen Zelltypen oder Zellzuständen, vergleichend auswerten und die einzelnen Komponenten durch **Massenspektrometrie** identifizieren.

Im Laufe der letzten Jahre sind zahlreiche Varianten der Kombination von Proteinauftrennung und Massenspektrometrie ausgearbeitet worden, welche nun einen Massendurchsatz ermöglichen.

Ebenso aufwändig gestaltet sich immer noch die Präparation von spezifischen Antikörpern, mit denen man Proteine identifizieren kann, oder die systematische Untersuchung von Wechselwirkungen zwischen Proteinen. Dennoch haben sich bereits viele Forschungslabors, auch in der pharmazeutischen Industrie, auf das Gebiet der Proteomics begeben, um sich speziellen biologischen und medizinischen Fragestellungen zu widmen, wie zum Beispiel der Identifizierung der während einer Krankheit veränderten Proteine, der Klassifizierung von Tumoren über ‚Proteinmarker‘ oder der Entwicklung neuer diagnostischer Methoden auf Proteinbasis [76, 77].

Es würde im Rahmen dieses Aufsatzes zu weit führen, alle heute etablierten Verfahren zu erläutern, die zur Isolierung, Auftrennung, Markierung, Identifizierung oder Klassifizierung von Proteinen dienen. Es bietet sich aber an, wichtige Fragestellungen zu berühren.

Proteine werden zur Wahrnehmung ihrer regulatorischen Aufgaben im Zellgeschehen verändert. Spezielle Enzyme werden mobilisiert, um Proteinfaktoren zu phosphorylieren oder zu dephosphorylieren; dadurch ändert sich deren Aktivität. Das Gleiche geschieht bei Methylierungen und Acetylierungen, die zum Beispiel helfen, die Histonproteine in der Chromatinstruktur der Gene zu verändern, um deren Zugänglichkeit zu modifizieren. Die chemische Kopplung von Proteinen mit bestimmten signal-gebenden Proteinen dient vielen Zwecken, zum Beispiel als ortsbestimmendes Etikett im zellulären Transport von Proteinen, als Signal für ihre Sekretion, oder als Etikett, das sie für den Abbau bestimmt.

Proteinkomplexe können in manchen Fällen viele verschiedene Proteine zusammenführen (manchmal über Hundert), die regulatorische Aufgaben oder Aufgaben im zellulären Transport von Komponenten übernehmen. Jedem Mitglied solcher Komplexe obliegt eine bestimmte Funktion, die für die Aktivität des Komplexes verantwortlich ist.

In ersten Ansätzen werden auch **Metabolome** von Organismen untersucht, weil man das Zusammenspiel aller Stoffwechselwege und aller Metabolite verstehen möchte.

Weiterhin interessieren die Forscher sich für die **Interactome** in Zellen, Geweben und ganzen Organismen. Als äußerst wichtig haben sich dabei die Wechselwirkungen (Interaktionen) zwischen DNA und Proteinen herausgestellt.

## **4.2 Wechselbeziehungen zwischen Genen und Proteinen**

Je mehr Daten zu Expressionsprofilen eines Organismus unter den verschiedensten Aspekten gesammelt werden, desto aufwändiger wird ihre Verarbeitung und Interpretation [z.B. 77]. Selbst bei

fortschreitender Einsicht können sie aber auch nur bedingt Auskunft darüber geben, in welcher Weise die Aktivitäten der einzelnen Gene miteinander vernetzt sind. Schließlich sind es nicht die Gene alleine, die in Wechselbeziehungen stehen, sondern die von ihnen codierten Proteine, und auf diese konzentrierte sich die Aufmerksamkeit. Die Herausforderung in der ‚Postgenom-Ära‘ bedeutete also, mehr darüber zu erfahren, wie die Aktivität von Genen im Zellgeschehen gesteuert wird, wie die Proteine einer Zelle oder eines Organismus funktionieren, wie sie interagieren. Wir wissen zwar schon länger, dass Gene grundsätzlich dadurch an- oder abgeschaltet werden können, dass spezielle **Regulatorproteine** (‚Transkriptionsfaktoren‘) an eigens dafür in den Genregionen vorgesehene DNA-Sequenzen (sogenannte Promotoren) binden, und dass es unter diesen Faktoren gewisse Hierarchien während der Entwicklung oder dem Wachstum eines Organismus gibt (Abbildung 25). Feinheiten im Zusammenspiel der Faktoren, die sich in vielen Fällen auch erst zu Proteinkomplexen zusammenfügen müssen, um aktiv zu sein, werden bisher nur exemplarisch verstanden; Abbildung 26 zeigt ein Beispiel.

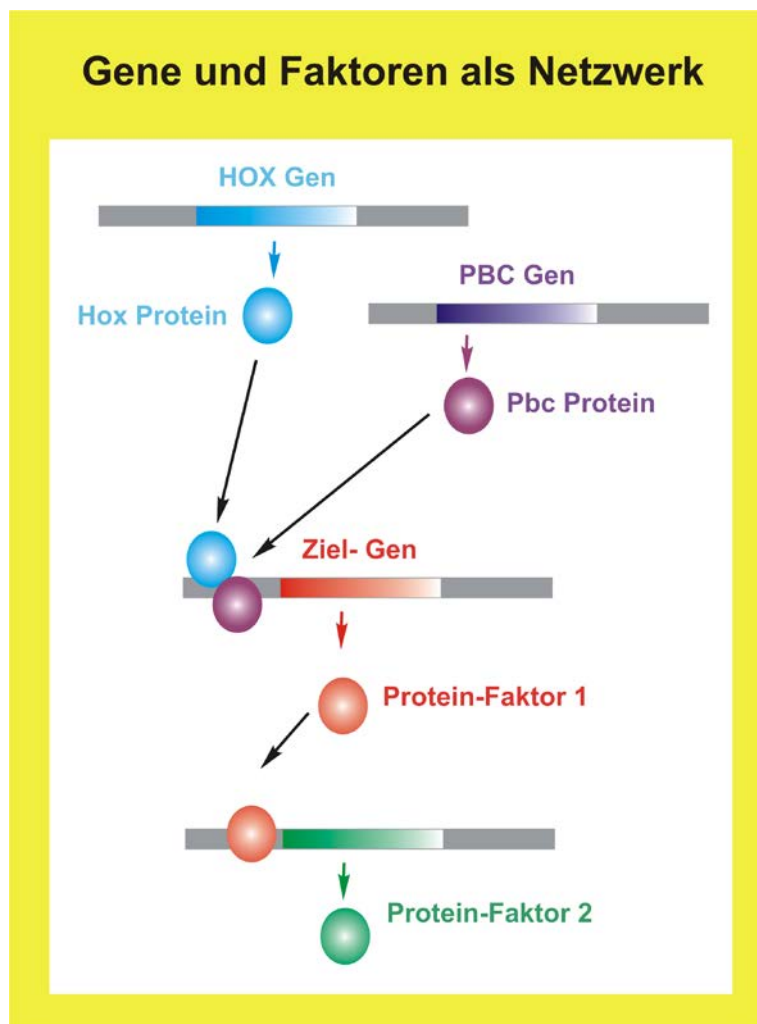


Abbildung 25: Genetisches Netzwerk am Beispiel eines HOX Gens.

HOX Gene (als Abkürzung für *homeobox Gene*) sind eine Gruppe von fast ubiquitär vorkommenden Regulatorgenen, die Hox Proteinfaktoren encodieren [78]. Beim Menschen wurden ca. 40 verschiedenen Varianten von HOX Genen gefunden. Sie wurden zuerst eingehend untersucht bei *Drosophila*, da sie dort die embryonale Entwicklung entlang der anterior-posterior (Kopf-Schwanz) Achse steuern: Nachdem sich die Embryonalsegmente gebildet haben, bestimmen Hox Proteine die Art des Segmentes für die Ausbildung der Beine, Antennen, oder Flügel. Bei Säugern determinieren Hox Proteine z.B. die

verschiedenen Rippen. Hox Proteine enthalten *Homeoboxen* und zeigen weitere abschnittsweise Übereinstimmungen. In vielen Säugern ist die Anordnung der HOX Gene die gleiche, sie folgt der Kopf-Schwanz Achse, und ihre Expression findet in dieser Reihenfolge statt. Man nennt das auch

„Kolinearitätsprinzip“. Homeobox-Proteine finden sich schon bei Hefen und bestimmen dort den „Geschlechtstyp“ (z.B. **a**-Zelle oder alpha-Zelle).

Jedes Hox Protein erfordert die Anwesenheit mindestens eines weiteren DNA-bindenden Partners, um den korrekten Satz von Zielgenen *in vivo* zu erkennen. Solches sind die sogenannten Pbc Proteine (*protein binding cofactors* - PBC): Sie bilden zusammen mit Hox-Proteinen kooperative DNA-bindende Komplexe und werden – vermutlich in Verbindung mit weiteren Cofaktoren - für die Steuerung der embryonalen Entwicklung durch die Hox-Proteine benötigt.

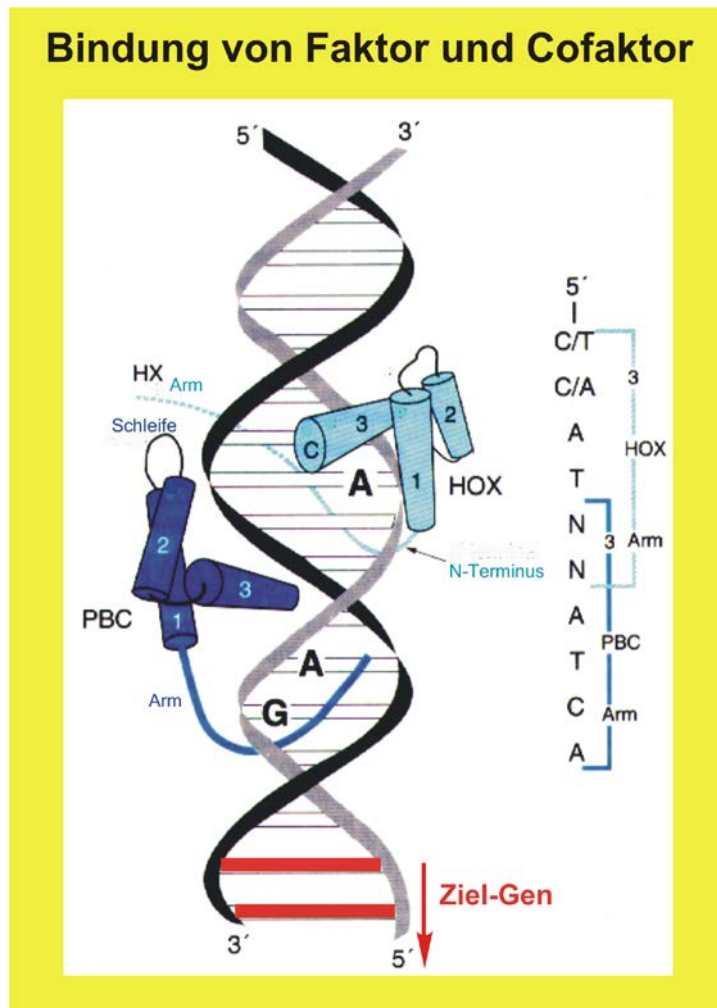


Abbildung 26: DNA-Bindung an die Regulatorregion eines Zielgens.

Hox Protein und Pbc Protein besitzen drei sog.  $\pm$ -Helices (schraubenartig gewickelte Abfolgen von Aminosäuren) und lange, nicht-strukturierte Arme. Die vom Hox Protein gebundene DNA-Abschnitt enthält meist die Sequenz T-A-A-T, gefolgt von zwei Basen (N-N), die mit unterscheiden, welches Homeobox-Protein binden soll. Diese Bindung allein ist jedoch noch zu unspezifisch (solche Sequenzen aus 6 Basen kommen statistisch zu häufig im Genom vor!), um gravierende Fehler in der Embryonalentwicklung zu vermeiden. Die Spezifität wird daher erhöht durch die (gleichzeitige) Bindung eines Cofaktors, in unserem Falle des Genproduktes des PBC Gens. Weitere solche Cofaktoren sind bekannt. Sie induzieren offenbar

Änderungen in der Konformation des Hox-Proteins, das erst anschließend als Regulator „wirksam“ wird.

#### 4.3 Nicht-codierende Gene im Genom

Zu den lange bekannten, nicht-codierenden Ribonukleinsäuren (RNA) gehören die ribosomalen RNAs (rRNA), Transfer-Ribonukleinsäuren (tRNAs), „small nuclear RNAs“ (snRNAs und snoRNAs), deren Gene unbedingt intakt sein müssen. Neuere Erkenntnisse weisen darauf hin, dass diese und alle codierenden Gene nicht allein den Genotyp eines Individuums ausmachen. Vielmehr wurden unzählige kleine, kurzlebige RNAs im gesamten Genom nebst ihren Genen identifiziert, welche die

Genexpression begleiten oder bestimmen (zeitlich und räumlich), oder sogar abschalten können, unter dem Sammelnamen kleine **ncRNAs** (small ,non-coding RNAs') zusammengefasst. Wegen der erkannten Wirkungsweise dieser kleinen RNAs werden sie heute in der Diagnostik und Therapie eingesetzt und dafür auch synthetisch hergestellt. Einen Überblick über die gefundenen Arten und Aufgaben dieser RNAs liefert eine Übersicht in der englischen Wikipedia [79]. Deutsche Beschreibungen gibt es in einer Ausgabe von *BIOspektrum* ([www.biospektrum.de](http://www.biospektrum.de)) aus dem Jahre 2009 [80 - 82].

#### 4.3.1 RNA-Interferenz (oder ,gene silencing')

Das Phänomen der **RNA-Interferenz** (RNAi) wurde zuerst bei Pflanzen entdeckt, hat sich aber als genetisches Regulativ in der gesamten Evolution erwiesen. Dafür sind kleine ncRNAs zuständig:

##### **'Small interfering RNA' (siRNA)**

In Metazoen (Vielzeller mit Ausnahme von Schwämmen) können siRNAs aus verschiedenen Vorläufern gebildet werden, u.a. aus langen doppelsträngigen RNA-Molekülen oder aus so genannter Antisense-RNA. siRNAs kommen im Kern und im Cytoplasma vor; sie übernehmen zum Beispiel die Genregulation auf der Translationsebene, indem sie mit messenger-RNAs (durch Basenpaarbildung) interagieren und diese zum Abbau veranlassen. Die siRNA wird in einen Komplex eingebunden, der Nukleasen enthält, welche dann die messenger-RNA im Innern spalten und weiter abbauen.

Möglicherweise gehören die siRNAs zu einem zellulären Verteidigungssystem (,innate immunity') gegen Viren, deren Genom aus doppelsträngiger DNA besteht.

##### **,'Micro-RNA' (miRNA)**

miRNAs sind kleine RNA-Moleküle (21-28 Nukleotide lang), die typischerweise komplementär zu messenger-RNAs in Metazoen sind. Bindung an die komplementären Sequenzen einer mRNA unterdrückt die Umsetzung der Botschaft in ein Protein; das befördert auch den Abbau der mRNA.

##### **,'Piwi-interacting RNA' (piRNA)**

piRNAs stellen die größte Klasse kleiner nicht-codierender RNA in tierischen Zellen dar, die sich aus einer großen Zahl von Vorläufer-Molekülen ableiten; das sind u.a. auch repetitive DNA und Transposons (vgl. Abschnitt 2.5.1; Abbildung 16). Ihre Biosynthese ist aber noch nicht gut verstanden; sie dürften auf verschiedenen Ebenen nach der Transkription, aber auch auf der Ebene der Chromatinstruktur, in die Zellregulation eingreifen.

#### 4.3.2 Andere nicht-codierende RNAs

##### **,'Antisense-RNA' (asRNA)**

asRNAs sind einzelsträngige RNAs, die komplementär zu messenger-RNAs sind und vom ,Gegenstrang' der codierenden DNA-Sequenz transkribiert werden; sie wurden schon 1978 entdeckt.

Die Transkription des Gegenstranges geschieht relativ häufig in der Zelle, jedoch sind – mit einigen Ausnahmen – die *in vivo* Funktionen der asRNAs weitgehend unbekannt.

Antisense-RNAs haben sich dagegen als wichtige Werkzeuge in der Erforschung von Genfunktionen erwiesen. Wird eine asRNA (13-28 Nucleotid lange Moleküle reichen aus) in eine Zelle eingeführt, so bindet sie an ihre ‚zugehörige‘ messenger-RNA und kleistert gewissermaßen diese zu, so dass infolge der sterischen Blockade keine Translation erfolgt. Verstärkt kann der Effekt dadurch werden, dass man statt asRNA eine asDNA verwendet und mit Ribonuklease H koppelt, so dass die messenger-RNA zu >80% abgebaut wird.

### **Ribozyme**

Ribozyme sind einzelsträngige, katalytische Ribonukleinsäuren, die in eine hammerartige Sekundärstruktur gefaltet und in der Lage sind, andere RNA-Moleküle an einer bestimmten Stelle zu schneiden, sobald sich komplementäre Sequenzen aneinander binden. Sie wurden 1989 entdeckt und finden sich natürlicherweise in einigen RNA-prozessierenden Enzymen der Zelle, aber auch in Viren. Der bekannte Mechanismus wird in der Forschung ausgenutzt, Ribozyme synthetisch zu erzeugen, welche die Expression ‚unerwünschter‘ Gene stilllegen. Solche könnten also als Waffen gegen Krankheitsgene eingesetzt werden.

### **„Long non-coding RNA (LncRNA)“**

Lange (>200 Nukleotide) nicht-codierende RNAs sind eine eher kürzliche Entdeckung durch das FANTOM Programm (Funktionale Annotation tierischer cDNA) [83]. Man fand etwa 35.000 solcher Transkripte von 10.000 verschiedenen Genorten des menschlichen Genoms; sie ähneln messenger RNAs, besitzen aber keine durchgehenden offenen Leserahmen. Insgesamt ist es schwierig, codierende Gene von solchen für nicht-codierende (LncRNAs) zu unterscheiden, aber es steht fest, dass LncRNAs fünfmal so häufig auftreten wie messenger-RNAs. Im Gegensatz zu den kleinen ncRNAs sind die LncRNAs in verschiedenen Organismen nicht evolutionär konserviert.

Offenbar greifen LncRNAs in die zelluläre Regulation auf den verschiedensten Ebenen ein. Sie werden gehandelt als Mediatoren: für epigenetische Prozesse (Umstellung der Chromatinstruktur oder Aktivierung von Entwicklungsgenen); für das so genannte Imprinting (Expression entweder der vom Vater oder der von der Mutter ererbten Gene); für die notwendige Inaktivierung eines der beiden X-Chromosomen bei weiblichen Nachkommen; für die Beteiligung am Alterungsprozess und an der Entstehung von Krankheiten. Ehe man weitere Erkenntnisse gesammelt hat, bleiben die Funktionen der LncRNAs weitgehend spekulativ.

## **4.4 Regulation der Zellaktivität**

Das menschliche Genom ist also wesentlich komplexer als man bisher angenommen hat. Je mehr man gräbt, desto unwahrscheinlicher wird es anzunehmen, dass der größte Teil des Genmaterials nur Ballast sei und keine Bedeutung für die Umsetzung der Information sowie die Regulation im Gesamtorganismus besitze. Unser vereinfachtes Bild darüber, wie die Aktivität einer Zelle gesteuert



wird (Abbildung 27), muss nach all den Entdeckungen der letzten Jahre wesentlich erweitert werden und neue Aspekte aufgreifen.



Abbildung 27: Die Aktivität einer Zelle (Zellwachstum, Zellteilung) kann durch äußere (exogene) und innere (endogene) Signale gesteuert werden. Äußere Signale sind Umweltfaktoren (z.B. Verfügbarkeit von Nährstoffen) oder über ‚Fernsteuerung‘ wirkende Faktoren (Hormone, Wachstumsfaktoren, von anderen Zellen abgegebenen Faktoren - hier Reize genannt -). Sie führen meist über ganze Kaskaden von Genen und ihren Genprodukten zur An- oder Abschaltung von Regulatorgenen im Zellkern. Differenzierung einer Zelle erfolgt meist ebenfalls durch An- oder Abschaltung von Regulatorgenen, die durch das genetische Programm angestoßen werden.

Ähnlich beschaffen wie das menschliche Genom sind die Genome der großen Menschenaffen. Man sagte, Mensch und Schimpanse (unser nächster Verwandter) unterschieden sich nur 1% in ihren Genen, bezogen auf Sequenzunterschiede, d.h. sie seien zu 99% genetisch identisch. (Sogar die Maus stimmt zu 98% in ihren Gensequenzen mit dem Menschen überein.) Diese Aussage gilt aber nur so lange, wie wir an den codierenden Genen ablesen konnten. Es gibt auf anderer Ebene deutliche Unterschiede: Schimpansen haben 24 statt 23 Chromosomen, weil das humane Chromosom 2 aus zwei kleineren Chromosomen (Schimpansen 2A und 2B) des Menschen zusammengewachsen ist. Die genetische Varianzbreite beträgt bei menschlichen Individuen durchschnittlich 0.5 Prozent, bei Schimpansen 2.7 Prozent, da letztere seit ihrer Trennung vom gemeinsamen Vorfahren vor 6 Mio Jahren viele Genduplikationen oder andere Veränderungen erfahren haben. Der Einfluss der oben geschilderten, zusätzlichen Faktoren (kurze und lange nicht-codierende RNAs) darf nicht vernachlässigt werden; Menschen und Schimpansen könnten sich hier gravierend unterscheiden. Alle genomischen Vergleiche helfen aber bisher nicht weiter, welche Besonderheiten den modernen Menschen hervorgebracht haben.

Eine Feststellung ist heute gesichert: *homo sapiens* hat sich in Afrika aus hominiden Vorfahren entwickelt und von dort aus vor etwa 300.000 – 200.000 Jahren über Afrika ausgedehnt und nach Asien und Europa verbreitet („Rassenbildung“); das geht aus paläontologischen Funden und Vergleichen der mitochondrialen DNA hervor. Besondere Erwartungen knüpfte man an die Sequenzierung des Genoms des Neandertalers durch Svante Pääbo im Leipziger Max-Planck-Institut für Anthropologie [84]. Inzwischen ist die Auswertung der Daten so weit gediehen, dass sich eine

Reihe von Fragen beantworten lassen: (1) Der Neandertaler ist mit uns heutigen Menschen verwandt, er entwickelte sich ebenfalls in Afrika vor etwa 600.000 Jahren, begann aber bereits ~300.000 Jahre früher als *homo sapiens* sich in Richtung Norden auszubreiten. (2) Bemerkenswert ist, dass die mitochondriale DNA des Neandertalers von der des modernen Menschen abweicht. Das bedeutet, wir haben keine mitochondriale DNA vom Neandertaler übernommen. (3) Neandertaler und moderner Mensch können durchaus längere Zeit nebeneinander (oder miteinander) gelebt haben. Dass sie Gene vom Genpool des jeweilig anderen übernommen haben, ist damit möglich, aber wenn, geschah dies nur in geringem Umfang. (4) Das Aussterben des Neandertalers vor ~30.000 Jahren ist belegt; die Ursache dafür bleibt unbekannt. Die ältesten Funde humaner Fossilien im Mittleren Osten sind 93.000 Jahre alt; die ältesten Funde von Neandertalern dort 60.000 Jahre. So kann 30.000 Jahre eine friedliche Koexistenz bestanden haben. (5) Neandertaler waren mit großer Wahrscheinlichkeit einer artikulierten Sprache mächtig. Das FOXP2- Protein, welches die Entwicklung des Gehirns und insbesondere die Artikulationsfähigkeit steuert, ist bei beiden Spezies, *Homo sapiens* und Neandertaler identisch und zeichnet sich durch den Austausch zweier Aminosäuren aus, die sonst nirgendwo bei Säugern, auch nicht bei den großen Menschenaffen, auftritt, sondern nur bei Menschen. Ist die ‚Menschwerdung‘ im Verlaufe der Evolution also durch den ‚Erwerb‘ besonderer kognitiver Fähigkeiten und durch gleichzeitige Entwicklung anatomischer Besonderheiten des Stimmapparates zu erklären?

## Ausblick

Die aus den Genomprojekten gewonnene Information wird auch in den kommenden Jahren exponentiell wachsen. Durch die Genomforschung haben wir - wenn auch bisher nur in begrenztem Umfang - Einblick in die Erbanlagen der unterschiedlichsten Lebewesen gewonnen und beginnen zu verstehen, wie sich im Verlaufe von etwa vier Milliarden Jahren in der Natur biologische Prozesse entwickelt haben, die in überaus komplexen Abläufen optimiert worden sind: Das „Darwin-Jahr“ 2009 lässt grüßen! Ein Novum aber ist, dass wir heute diese genetische Information nutzen können. In der neuen industriellen Revolution steht nicht die Gewinnung und Verarbeitung von Rohstoffen, sondern die Beherrschung der Information im Mittelpunkt: Hier berühren sich die Informationstechnologie, Computer und Internet, mit der Genomforschung aufs engste. Unverkennbar hat die Notwendigkeit zur Einführung von ‚Nano-Techniken‘ und Massendurchsatz in der Molekularbiologie ihre Parallele in der Entwicklung der Mikroelektronik und der Computertechnik. Wie bei der letzteren wird es auch in der Bioinformatik immer wichtiger werden, intelligente Systeme zu entwickeln, die biologisch relevante Muster in massiven Datenmengen erkennen können und geeignet sind, Wechselwirkungen zwischen Biomolekülen vorherzusagen, unter anderem eine unerlässliche Hilfe zum Entwurf neuer Therapeutika und Diagnostika.

## Literatur

### Allgemein

Ulrich Schäfer (Herausgeber), Bruce Alberts (Autor), Alexander Johnson (Autor), Julian Lewis (Autor), Martin Raff (Autor), Keith Roberts (Autor), Peter Walter (Autor), Bärbel Häcker (Übersetzer), Claudia Horstmann (Übersetzer) **Molekularbiologie der Zelle**, 5. Auflage, April 2011, VCH Verlagsgesellschaft, Weinheim.

### Spezielle Zitate

- [1] A Revolution in Biology (1980) Sonderheft September 1980, *Science*, **209**.
- [2] D.A. Jackson, R.H. Symons and P. Berg (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 2904-2909.
- [3] P. Berg *et al.* (1974) Potential Biohazards of Recombinant DNA Molecules. *Science*, **185**, 303.
- [4] P. Berg, D. Baltimore, S. Brenner, R.O. Roblin III, M.F. Singer (1975) Asilomar Conference on Recombinant DNA Molecules. *Science*, **188**, 991-994.
- [5] F. Sanger, S. Nicklen and A.R. Coulson (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-1. <http://www.theguardian.com/science/2013/nov/20/frederick-sanger>
- [6] Quail, Michael; Smith, Miriam E; Coupland, Paul *et al.* (1 January 2012). "A tale of three next generation sequencing platforms: comparison of Ion torrent, pacific biosciences and illumina MiSeq sequencers". *BMC Genomics* **13** (1): 341. doi:10.1186/1471-2164-13-341. PMC 3431227. PMID 22827831.
- [7] Stoddart, D; Heron, AJ; Mikhailova, E; Maglia, G; Bayley, H (2009 May 12). "Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore.". *Proceedings of the National Academy of Sciences of the United States of America*, **106** (19): 7702–7. Bibcode:2009PNAS..106.7702S. doi:10.1073/pnas.0901054106. PMC 2683137. PMID 19380741.
- [8] Massimiliano Di Ventra (2013). "Fast DNA sequencing by electrical means inches closer". *Nanotechnology*, **24**, 342501
- [9] Ohshiro T *et al* (2012) *Sci. Rep.*, **2**, 501–2 507
- [10] <http://en.wikipedia.org/wiki/PCR>
- [11] Sequencing industry  
<http://en.wikipedia.org/w/index.php?title=Special%3ASearch&profile=default&search=Sequencing+industry&fulltext=Search>
- [12] The yeast genome directory (29 May 1997) *Nature* 387, issue 6632S
- [13] A. Goffeau, B.G. Barrell, H. Bussey, R.W. Davis, B. Dujon, H. Feldmann, F. Galibert, J.D. Hoheisel, C. Jacq, M. Johnston, E.J. Louis, H.W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin and S.G. Oliver (1996) Life with 6000 genes. *Science* **274**, 546, 563-1.
- [14] Horst Feldmann: Yeast - Molecular and Cell Biology. Second edition 2012, WILEY Publishers
- [15] Exploration of the Hemiascomycetous Yeasts, ed. H. Feldmann (2000). *FEBS Letters*, 487, vol.1.
- [16] Dujon, B. (2010) Yeast evolutionary genomics. *Nature Reviews Genetics*, **11**, 512-524.
- [17] Vormals 'The Institute of Genome Research' (TIGR), jetzt 'Craig Venter Institute' (JCVI):  
<http://www.jcvi.org/cms/research/projects/tdb/overview/>
- [18] National Institute of Health, Bethesda (NIH): <http://www.ncbi.nlm.nih.gov/>
- [19] The *C. elegans* Sequencing Consortium (1998) *Science*, **282**, 2013.
- [20] C.I. Bargmann *et al.* (1998) Signal transduction in the *Caenorhabditis elegans* nervous system. *Science*, **282**, 2029.
- [21] G.M. Rubin *et al.* (2000) The Drosophila genome: comparative genomics of eukaryotes, *Science*, **287**, 2204-2215.
- [22] NIH: <http://en.wikipedia.org/wiki/NIH>; Ethische Aspekte des Humangenom-Projektes:  
<http://www.sciencemag.org/content/265/5181/2035.long>
- [23] Humangenom Organisation (HUGO). <http://www.hugo-international.org/>
- [24] <http://www.genenames.org/>
- [25] F.S. Collins *et al.* (1998) New goals for the U.S. Human Genome Project. *Science*, **282**, 682-689.
- [26] I. Dunham, N. Shimizu, B.A. Roe *et al.* (1999) *Nature*, **402**, 489-495.
- [27] M. Hatori *et al.* (2000) The DNA sequence of human chromosome 21, *Nature*, **405**, 311-319
- [28] International Human Genome Sequencing Consortium (2001), *Nature*, **409**, 860-921.
- [29] J.C. Venter *et al.*, (2001) *Science*, **291**, 1304-1351.
- [30] R.H. Waterston, E.S. Lander, J.E. Sulston (2002) On the sequencing of the human genome, *Proc. Natl. Acad. Sci. USA*, **99**, 3712-3716.

- [31] Gregory SG, Barlow KF, McLay KE et al. (May 2006). "The DNA sequence and biological annotation of human chromosome 1". *Nature* **441** (7091): 315–21. [Bibcode:2006Natur.441..315G](#). [doi:10.1038/nature04727](#). PMID 16710414. Vgl. auch ref. 45.
- [32] Cancer Genome Anatomy Project: <http://www.ncbi.nlm.nih.gov/gquery/?term=Cancer+Genome+Anatomy+Project>  
 Navigiere zu **1507986 EST**: expressed sequence tag sequences
- [33] Human Genome Diversity Project: E. Pennisi et al. (1997) NRC OKs long-delayed survey of human genome diversity. *Science*, **278**, 568.
- [34] Gardner MJ, Hall N, Fung E et al. (October 2002). "Genome sequence of the human malaria parasite *Plasmodium falciparum*". *Nature* **419** (6906): 498–511. [Bibcode:2002Natur.419..498G](#). [doi:10.1038/nature01097](#). PMID 12368864.
- [35] Holt RA, Subramanian GM, Halpern A et al. (October 2002). "The genome sequence of the malaria mosquito *Anopheles gambiae*". *Science* **298** (5591): 129–49. [Bibcode:2002Sci...298..129H](#). [doi:10.1126/science.1076181](#). PMID 12364791.H
- [36] Malaria: <http://en.wikipedia.org/wiki/Malaria#Research>
- [37] Berriman M, Ghedin E, Hertz-Fowler C et al. (July 2005). "The genome of the African trypanosome *Trypanosoma brucei*". *Science* **309** (5733): 416–22. [Bibcode:2005Sci...309..416B](#). [doi:10.1126/science.1112642](#). PMID 16020726.
- [38] Schlafkrankheit: [http://en.wikipedia.org/wiki/African\\_trypanosomiasis](http://en.wikipedia.org/wiki/African_trypanosomiasis)
- [39] Peacock CS, Seeger K, Harris D et al. (July 2007). "Comparative genomic analysis of three *Leishmania* species that cause diverse human disease". *Nature Genetics* **39** (7): 839–47. [doi:10.1038/ng2053](#). [PMC 2592530](#). PMID 17572675.
- [40] Leishmaniose: <http://en.wikipedia.org/wiki/Leishmaniasis>
- [41] Analysis of the genome of the flowering plant *Arabidopsis thaliana* (2000), *Nature*, **408**, 796-815.
- [42] Jun Yu et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica) *Science*, **296**, 79-91.
- [43] S.A. Goff et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica) *Science*, **296**, 92-100.
- [44] Pflanzengenome: [http://en.wikipedia.org/wiki/List\\_of\\_sequenced\\_plant\\_genomes](http://en.wikipedia.org/wiki/List_of_sequenced_plant_genomes)
- [45] [http://en.wikipedia.org/wiki/List\\_of\\_sequenced\\_genomes](http://en.wikipedia.org/wiki/List_of_sequenced_genomes) (Überblick sequenzierter Genome)
- [46] Methanococcus: Smith, DR (1997). "Complete genome sequence of *Methanobacterium thermoautotrophicum* deltaH: functional analysis and comparative genomics". *The Journal of Bacteriology* **179** (22): 7135–55. [PMC 179657](#). PMID 9371463.
- [47] <http://www.jcvi.org/cms/research/groups/> (Neuere Genomprojekte)
- [48] [http://www.jcvi.org/cms/index.php?id=629&no\\_cache=1&tx\\_ttnews\[tt\\_news\]=3284/](http://www.jcvi.org/cms/index.php?id=629&no_cache=1&tx_ttnews[tt_news]=3284/) (Human Mikrobiom)
- [49] <http://www.jcvi.org/cms/research/projects/gos/overview/> (Kleinlebewesen der Weltmeere)
- [50] <http://www.jcvi.org/cms/research/projects/huref/overview/> (Diploides Genom)
- [51] <http://www.vci.de/dib/> (Übersichten zur Biotechnologie);  
 USA: [http://en.wikipedia.org/wiki/National\\_Center\\_for\\_Biotechnology\\_Information](http://en.wikipedia.org/wiki/National_Center_for_Biotechnology_Information)
- [52] <http://www.vfa.de/download/broschuere-biopharmazeutika.pdf> (Pharmaforschung)
- [53] [http://www.nature.com/nbt/journal/v28/n9/fig\\_tab/nbt0910-917\\_T1.html](http://www.nature.com/nbt/journal/v28/n9/fig_tab/nbt0910-917_T1.html) (Biotherapeutika)
- [54] <http://en.wikipedia.org/wiki/GMO> (Gentechnisch veränderte Organismen)
- [55] Murray, Joe (2010). Genetically modified animals. Canada: Brainwaving.  
<http://www.brainwaving.com/2010/07/28/genetically-modified-animals/>
- [56] Andrew Pollack for the New York Times. "An Entrepreneur Bankrolls a Genetically Engineered Salmon"  
 Published: May 21, 2012. Accessed October 7, 2012
- [57] <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1878528/>
- [58] Mukoviszidose: [http://en.wikipedia.org/wiki/Cystic\\_fibrosis](http://en.wikipedia.org/wiki/Cystic_fibrosis)
- [59] <http://www.the-scientist.com/?articles.view/articleNo/33166/title/Gene-Therapy-Arrives-in-Europe/>
- [60] Grüne Gentechnik: [http://de.wikipedia.org/wiki/Transgene\\_Pflanzen](http://de.wikipedia.org/wiki/Transgene_Pflanzen)
- [61] Ethische Aspekte: [http://de.wikipedia.org/wiki/Transgene\\_Pflanzen#Ethische\\_Aspekte](http://de.wikipedia.org/wiki/Transgene_Pflanzen#Ethische_Aspekte)
- [62] Gentechnikgesetz: [http://de.wikipedia.org/wiki/Gentechnikgesetz\\_\(Deutschland\)](http://de.wikipedia.org/wiki/Gentechnikgesetz_(Deutschland))
- [63] [http://de.wikipedia.org/wiki/Richtlinie\\_2001/18/EG\\_\(Freisetzungsrichtlinie\)](http://de.wikipedia.org/wiki/Richtlinie_2001/18/EG_(Freisetzungsrichtlinie))
- [64] Weiße Biotechnik: [http://de.wikipedia.org/wiki/Industrielle\\_Biotechnologie](http://de.wikipedia.org/wiki/Industrielle_Biotechnologie)
- [65] Nachwachsende Rohstoffe: <http://www.bmbf.de/de/6955.php>
- [66] <http://www.biotechnologie.de/>
- [67] <http://de.wikipedia.org/wiki/Biotreibstoff>
- [68] <http://en.wikipedia.org/wiki/Biofuel>
- [69] <http://en.wikipedia.org/wiki/E.coli>

- [70] <http://news.discovery.com/tech/alternative-power-sources/bacteria-powered-light-bulb-is-electricity-free-130815.htm>
- [71] z.B.: J.L. DeRisi *et al.* (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science*, **278**, 680-686.
- [72] E.A. Winzeler, D.D. Shoemaker, A. Astromoff *et al.* (1999) *Science*, **285**, 901-906.
- [73] Bulyk, M.L. (2006) DNA microarray technologies for measuring protein-DNA interactions. *Current opinion in Biotechnology*, **17**, 422-430
- [74] Chen, R. and Snyder, M. (2010) Yeast proteomics and protein microarrays. *Journal of Proteomics*, **73**, 2147-2157
- [75] Costanzo, M. *et al.* ((2010) The genetic landscape of a cell. *Science*, **327**, 425-431
- [76] S. Nock, P. Wagner (2000) Proteomics: Die post-genomische Revolution. *Chemie in unserer Zeit* **34**, 348-354.
- [77] Supplement to Trends in Biotechnology (2001) *Trends in Proteomics*, **Vol.19**, No.10
- [78] HOX [http://en.wikipedia.org/wiki/Hox\\_genes](http://en.wikipedia.org/wiki/Hox_genes)
- [79] Übersicht ncRNAs:  
<http://en.wikipedia.org/w/index.php?search=Therapy+with+small+RNA&title=Special%3ASearch>
- [80] G. Stoecklin (2009) mRNA Abbau im Zytoplasma. *BIOspektrum* **15** (5), 505-508.
- [81] J. Hackermüller, F. Horn, P.F. Stadler, B. Buller (2009) Die neue Welt der ncRNAs und ihre medizinische Relevanz. *BIOspektrum* **15** (5), 509-511.
- [82] C. Becker, I. Riedmeier, M.W. Paffl (2009) RNA-Qualitätskontrolle in der Genexpressions-Analytik. *BIOspektrum* **15** (5), 512-515.
- [83] LcnRNAs: <http://en.wikipedia.org/wiki/LncRNA>
- [84] Neandertaler: <http://edge.org/conversation/mapping-the-neanderthal-genome>

## Glossar

Allel	Bezeichnet eine doppelt angelegte genetische Information.
Archea	(Archaeen; Archebakterien) Einzellige Organismen (⇒Prokaryonten) mit einem meist in sich geschlossenen DNA-Molekül (auch als zirkuläres Chromosom bezeichnet), das ein Kernäquivalent darstellt. Sie bilden eine eigene (dritte) systematische Domäne neben den Eubakterien und den ⇒Eukaryoten. Viele Arten der Archaeen sind an extreme Milieubedingungen angepasst, gehören damit zu den ⇒Extremophilen.
Autosomen	Bezeichnung für alle ⇒Chromosomen, die nicht geschlechts-bestimmende Information enthalten.
Bakterien	(Eubakterien) Diese bilden eine eigene Domäne unter den drei Domänen der Lebewesen. Es sind ⇒Prokaryonten, die keinen Kern und keine Mitochondrien besitzen, aber extrachromosomale Elemente wie Plasmide aufnehmen können. Wie ⇒Eukaryoten können Bakterien von Viren infiziert werden, die als Phagen (oder Bakteriophagen) bezeichnet werden. Vorläufer von Bakterien sind als Symbionten in Eukaryoten als ⇒Mitochondrien übernommen worden.
Bänderung	In Chromosomen aus höheren Organismen lassen sich durch Färbung mit verschiedenen Agenzien Bandmuster erzeugen. Färbung kann auch euchromatische Regionen (⇒ Euchromatin) von heterochromatischen (⇒ Heterochromatin) unterscheiden.
Basentriplett	Drei nebeneinanderliegende Basen werden als Basentriplett bezeichnet. Sie fungieren als Codons im Genetischen Code. Aus den vier Basen können 64 ⇒Codons ( $4^3$ Kombinationen mit Wiederholung) gebildet werden.
Chromosom	Eine Einheit, in der viele Gene zusammengefasst sind. Bei höheren Organismen ist die (doppelsträngige) DNA als ⇒Chromatin ‚verpackt‘. Die DNA ist abschnittsweise (~146 Basenpaare) um einen Kern aus vier basischen Proteinen (⇒Histone) gewickelt, das sog. ⇒Nukleosom. Diese ‚Perlschnur‘ wird durch ein weiteres Histon in der Packung verdichtet. Die genetische Information kann nur abgerufen werden, wenn sich die Chromatinstruktur durch eine geregelte Modifizierung der Histone entfaltet. Auch bei Bakterien spricht man vom ‚Chromosom‘, das meist eine ringförmig geschlossene (doppelsträngige) DNA mit stabilisierenden Proteinen darstellt.
Codon	Bezeichnung für ein Basentriplett, das bei der ⇒Translation einer bestimmten Aminosäure zugeordnet ist. Im Standardcode fungieren drei Codons als Stoppsignale (UUA, UAG, und UGA).
Cyanobakterien	Neuere Bezeichnung für blau-grüne Algen.
dicotyl	zweikeimblättrig
DNA-Polymerase	Sammelbegriff für Enzyme, die an der Duplikation von DNA-Molekülen beteiligt sind. Für die identische Reduplikation der menschlichen DNA sind fünf DNA-Polymerasen mit unterschiedlichen Funktionen erforderlich.
Dominant	In der Genetik wird zwischen dominanten und rezessiven ⇒Allelen unterschieden. Dominante Allele setzen sich bei der Merkmalsausprägung über die rezessiven durch.
Enzym	Protein, welches Stoffumsetzungen katalysiert.
<i>Escherichia coli</i>	( <i>E. coli</i> ) Bezeichnung für eine Reihe von endogenen Bakterien. Nicht-pathogene Stämme von <i>E.coli</i> besiedeln die menschliche Darmflora und helfen bei Verdauungsvorgängen. Infektionen mit pathogenen Stämmen (die ⇒Toxine ausschütten) führen zu unliebsamen Erkrankungen. <i>E. coli</i> werden als prokaryotisches Modell und in der Biotechnik als ‚Bioreaktoren‘ eingesetzt.
Euchromatin	Euchromatin besteht aus Chromosomenabschnitten, die transkribierbare genetische Information enthalten. Bei Färbung zeigt sich eine ⇒Bänderung.
Eukaryot	Organismus, dessen zelluläre DNA in einem, von einer Membran umschlossenen, Kern untergebracht ist. Eukaryoten besitzen weitere Zellorganellen (siehe auch ⇒ Mitochondrien).
Exon	Codierende Abschnitte in Vorstufen von ⇒messenger-RNA, die von ⇒Introns (nicht-codierende Abschnitte) unterbrochen werden.
Extremophile	Bezeichnung für Organismen, die unter extremen Milieubedingungen leben können oder leben müssen (z.B. hoher Druck, hohe oder niedrige Temperatur, hoher Salzgehalt, spezielle pH-Werte des Mediums...)
FISH	Kurz für Fluoreszenz in situ Hybridisierung, eine zytogenetische Technik, entwickelt für biomedizinische Anwendungen. Sie wird benutzt, um die An- oder Abwesenheit bestimmter DNA-Sequenzen auf Chromosomen zu sehen: Mit fluoreszenten Farbstoffen versehene ‚Proben‘ heften sich durch Hybridisierung (Basenpaarung) an entsprechende Sequenzen in der DNA. Beobachtet können diese Stellen durch Fluoreszenz-Mikroskopie werden.



Gelelektrophorese	Verfahren zur Auftrennung von Proteinen oder Nukleinsäuren. Als Träger werden Gele aus natürlichen (z.B. Agarose) oder synthetischen (z.B. Acrylamid) Polymeren verwendet; das zunächst flüssige Material wird heute meist zwischen zwei rechteckige Platten gegossen (bis zu 20 cm Breite und 80 cm Länge), die nur wenige Millimeter Abstand voneinander haben. In der eindimensionalen G. werden auf der Oberseite des Gels Taschen vorgeformt, in die nach Polymerisation die zu trennenden Proben eingetragen werden. An die Anordnung wird ein Hochspannungsfeld angelegt, wodurch sich die aufgetragenen Proben in ‚Laufrichtung‘ in ihre Komponenten auftrennen. Dabei spielt die durchschnittliche Ladung der zu trennenden Komponenten eine maßgebliche Rolle, die auch die Stromrichtung und den pH-Wert des Lauf-Puffers bestimmt. In der zweidimensionalen G. wird das Gel nach dem ersten Lauf meist um 90° gedreht und unter anderem pH-Wert nochmals ins Hochspannungsfeld versetzt. Die aufgetrennten Komponenten ergeben dann punktförmige Muster. Es sind zahlreiche Vorschriften und vorgefertigte Gele im Handel. Die Gelelektrophorese einer einzelnen Probe kann auch in einem zylindrischen Gefäß (bis zu Kapillardicke) erfolgen.
Gen	DNA-Abschnitt in einem Genom, das als Genprodukt entweder ein Protein oder eine nicht-codierende Sequenz (⇒ncRNA) beinhaltet.
Genfamilie	Bezeichnung von Genen, die in der Funktion gleiche oder ähnliche Genprodukte liefern.
Genkarte	Übersicht über alle Genorte (Genloci), die in einem Genom identifiziert werden.
Genom	Die Gesamtheit aller in einem Organismus enthaltenen Gene. Bei ⇒Viren spricht man ebenfalls vom Genom als Träger der gesamten genetischen Information.
Genort	Auch Locus, meint einen physikalisch oder genetisch ermittelten Abschnitt der genetischen Information, in dem bestimmte Funktionen lokalisiert sind
Geschlechtschromosomen	Neben den Autosomen gibt es spezielle Chromosomen, die geschlechts-spezifische Merkmale vermitteln. Beim Menschen wird das ‚weibliche‘ Chromosom als X-Chromosom, das ‚männliche‘ als Y-Chromosom bezeichnet.
Hämatopoese	Vorgang der Bildung aller Blutzellen aus einer hämatopoetischen ⇒Stammzelle.
Hefen	Unterordnung der Pilze. Heute sind etwa 3000 Arten von Hefen identifiziert und phylogenetisch eingeordnet.
Heterochromatin	Heterochromatin besteht aus sehr dicht gepacktem Chromatin (vgl. ⇒Chromosom)
Immunglobulin	Allgemeine Bezeichnung für Antikörper, die in verschiedene Typen eingeteilt werden.
Intron	Nicht-codierende Abschnitte in Vorstufen von ⇒messenger-RNA.
Karzinom	Tumorgewebe, das sich aus Epithelzellen ableitet, deren Genom so verändert oder beschädigt wurde, dass die Zellen abnormales Wachstum beginnen.
Klon	Organismen, die alle durch den Besitz identischer genetischer Information gekennzeichnet sind.
Kreuzung	1. Zuchtverfahren, in welchem Individuen von Tieren oder Pflanzen miteinander so lange gepaart werden, bis eine gewünschte Eigenschaft sich bei einem Nachkommen offenbart. 2. Analytisches Verfahren, um herauszufinden, welche Merkmale in welcher Weise auf die Nachkommen (F1-, F2-Generation; Rückkreuzungen) vererbt werden. 3. Analytisches Verfahren, vor allem bei Mikroorganismen, um herauszufinden, welchen Abstand (genetisch oder physikalisch) Gene voneinander haben (‚Kopplungsgrad‘).
Messenger-RNA	Bezeichnet die durch ⇒Transkription und anschließendes Prozessieren von Vorstufen entstandene RNA, die in der ⇒Translation benutzt werden soll.
Mikroorganismen	Kleinlebewesen, das sind meist einzellige Organismen, die in allen drei Domänen des Lebens vorkommen.
Mitochondrien	Zellorganellen, deren Vorkommen auf ⇒Eukaryoten beschränkt ist. Ihre Hauptaufgabe ist die Bereitstellung von Energie zur Synthese von zelleigenen Komponenten. Die Energie wird aus dem Abbau von Nahrungsstoffen (Zuckern, Fetten oder Proteinen) gewonnen; dazu ist der Austausch von Komponenten zwischen Cytoplasma und Mitochondrien erforderlich. Mitochondrien haben sich aus adaptierten Vorläufer-Bakterien entwickelt und verfügen über eine eigene genetische Information und einen eigenen Apparat zur Proteinbiosynthese.
monocotyl	einkeimblättrig
Mutation	Veränderung einer genetischen Information durch beabsichtigte oder unbeabsichtigte Manipulationen. Die Veränderungen können durch chemische Agenzien, Strahlung, etc. oder durch Fehlleistungen in der ⇒Replikation ausgelöst werden. Punktmutationen betreffen den Austausch oder die Auslöschung einer einzelnen Base, welche zu Unstimmigkeiten, z.B. in einem

	offenen Leseraster, führt. Deletionen bedeuten den Verlust (meist) längerer DNA-Abschnitte. Insertionen bedeuten den Zuwachs an DNA-Sequenzen. Mutationen führen letztlich zu Abweichungen der normalen Genexpression. (Eine Ausnahme ist der Basenaustausch innerhalb einer $\Rightarrow$ Codongruppe).
ncRNA	Nicht-kodierende Ribonukleinsäuren. Ihre Transkription liefert u.a. ‚Gebrauchs-RNAs‘ wie ribosomale RNAs ( $\Rightarrow$ rRNA); Transfer-RNAs ( $\Rightarrow$ tRNAs), kleine nukleäre RNA ( $\Rightarrow$ snRNAs) und kleine nukleoläre RNAs ( $\Rightarrow$ snoRNAs). Von Wichtigkeit, vor allem bei höheren Organismen sind kurze ncRNAs (miRNA, siRNA, piRNA), die zum Phänomen der RNA-Interferenz (oder auch ‚gene silencing‘) führen. Lange ncRNAs machen bei höheren Organismen den Hauptanteil transkribierter, aber nicht-codierender, Gene aus; sie greifen auf vielen Ebenen in die Regulation der Genexpression ein.
ORF	‚Open reading frame‘ (offener Leserahmen).
Parasit	Schädling, der Tiere oder Pflanzen befällt und (meist schwerwiegende) Krankheiten hervorruft
Pathogen	Als Adjektiv und Substantiv gebraucht, bedeutet eine krankmachende Eigenschaft bzw. einen krankmachenden Keim.
Photosynthese	Fähigkeit eines Organismus, Zucker aus Kohlendioxid und Wasser mithilfe von Sonnenenergie zu synthetisieren. Die damit befassten Organellen (Chloroplasten) versammeln alle notwendigen Komponenten. Zur Photosynthese sind neben allen grünen Pflanzen auch Cyanobakterien befähigt.
Pilze	(Fungi) sind eukaryotische Lebewesen, deren Zellen Mitochondrien und ein Zellskelett enthalten. In der biologischen Klassifikation bilden sie neben Tieren und Pflanzen ein eigenständiges Reich, zu dem sowohl Einzeller ( $\Rightarrow$ Hefen) als auch Vielzeller wie die Schimmelpilze und die Ständerpilze gehören. Der Artenreichtum wird auf 700.000 Spezies geschätzt.
Plasmid	Extrachromosomales (meist ringförmiges) DNA-Molekül, das ‚vorteilhafte‘ genetische Information angesammelt hat. In (pathogenen) Bakterien verleihen Plasmide zum Beispiel Resistenzen gegen bestimmte Antibiotika. Modifizierte Plasmide werden als Vektoren in der Einschleusung fremder DNA in Bakterien benutzt.
Polytänisierung	Endoreplikation (auch Polyploidisierung) ereignet sich in somatischen Geweben durch Programmänderungen des $\Rightarrow$ Zellzyklus. Die DNA repliziert, dagegen teilen sich der Zellkern und die Zelle nicht, werden aber größer. Den Vorgang nennt man Endoreduplikation, wenn der DNA-Gehalt jeweils exakt verdoppelt wird (bis zu $2^{12}$ Kopien). Endoreplikation kommt in der normalen Entwicklung bei Eukaryoten vor: bei Einzellern, bei Pflanzen (bei Nutzpflanzen fast obligat) und bei Tieren.
Prokaryot	Organismus, dessen zelluläre DNA <u>nicht</u> in einem, von einer Membran umschlossenen, Kern untergebracht ist. Die DNA liegt meist als zirkuläres Chromosom vor.
Pseudogen	Gen, welches keinen oder unterbrochene Leserahmen aufweist.
Radioautogramm	Mit Radioisotopen (z.B. $[^{32}\text{P}]$ -Phosphat oder $[^{33}\text{P}]$ -Phosphat) markierte Nukleinsäuren können durch Auflegen von Röntgenfilm auf den Träger, auf den sie aufgebracht wurden, sichtbar gemacht werden. Proteine können durch Markierung mit Jod-Isotopen radioautografiert werden.
Repetitives Element	Häufig im Genom vorkommendes DANN-Element, das sich über verschiedene Mechanismen darin verbreitet hat. Man spricht von $\Rightarrow$ Transposons und verschiedenen Typen von $\Rightarrow$ Retrotransposons.
Replikation	Bezeichnung für den Vorgang, in welchem aus der DNA eines Individuums eine zweite identische Kopie erzeugt wird.
Resistenz	Erworbener Widerstand gegen die Wirkung eines Antibiotikums oder eines anderen chemischen Agens. Bei Bakterien werden Resistenzgene gegenseitig durch $\Rightarrow$ Plasmide übertragen und können diese oft gegen zahlreiche Agenzien resistent machen (Streptokokkenresistenz; Hospitalismus).
Restriktion	Bakterien wehren sich gegen das Eindringen fremder DNA durch Enzyme (Restriktionsenzyme), welche die eingedrungene DNA zerschneiden, die zelleigene aber unversehrt lassen.
Retrotransposon	Transposon, dessen genetische Information eine Vorschrift zur Umwandlung der RNA in DNA enthält, welche ins Wirtsgenom integriert werden kann. Außerdem enthalten viele Retrotransposons Gene zur Herstellung von Kapselproteinen, zur Integration der DNA ins Wirtsgenom und oft auch mitgeschleppte Onkogene. Retrotransposons haben Ähnlichkeit mit $\Rightarrow$ Retroviren.
Retroviren	Eine große Familie von Viren mit einzelsträngiger RNA als genetischem Material, eingehüllt von Proteinkapseln. Im Gegensatz zu anderen RNA-Viren muss die RNA von Retroviren zunächst in DNA zurückgeschrieben werden (Retrotranskription), die dann in das Wirtsgenom integriert und wirksam werden

	kann. Retroviren infizieren vorrangig tierische Zellen, sind weit verbreitet und meist sehr wirtsspezifisch. Sie induzieren weit verbreitete Infektionskrankheiten. Beim Menschen wurden bisher fünf Retroviren identifiziert: HTLV-1 und HTLV-2 (Humanes T-lymphotropes Virus 1 bzw.2); HIV-I und HIV-II (Humanes Immunodefizienz-Virus, Typl bzw. TyplI); XRMV (Xenotropic murine leukemia virus-related virus).
Rezessiv	In der Genetik wird zwischen dominanten und rezessiven $\Rightarrow$ Allelen unterschieden. Dominante Allele setzen sich bei der Merkmalsausprägung über die rezessiven durch.
rRNA	Ribosomale RNA. Mehrere dieser RNAs mit unterschiedlicher Kettenlänge bilden das ‚Gerüst‘ der Ribosomen und dienen zusammen mit mehr als hundert ribosomalen Proteinen der Proteinbiosynthese.
Sequenzierung	Ermittlung der Art und Reihenfolge der Bausteine in informations-tragenden Polymeren, wie Peptiden, Proteinen, RNA und DNA, durch speziell entwickelte Techniken.
snoRNA	Small nucleolar RNAs. Etwa hundert solcher snoRNAs werden für die Prozessierung der Vorstufen der ribosomalen RNA benötigt und begleiten auch die Komplexbildung aus $\Rightarrow$ rRNAs und Proteinen in der Biosynthese der Ribosomen.
snRNA	Small nuclear RNAs sind maßgeblich an der Prozessierung von messenger-RNAs beteiligt, welche aus Exons (codierende Abschnitte) und Introns (nicht-codierende Abschnitte) bestehen. Die codierenden Abschnitte werden durch das sog. <i>spliceosome</i> zusammengefügt, einem Komplex aus hundert Proteinfaktoren und sechs snRNAs.
Spleißen	Vorgang, bei dem Vorstufen einer messenger-RNA so prozessiert werden, dass $\Rightarrow$ Introns herausgenommen und $\Rightarrow$ Exons aneinander gekettet (gespleißt) werden. Diesen Prozess vollführen Komplexe (sog. Spliceosomen) aus Proteinfaktoren und kleinen RNAs ( $\Rightarrow$ snRNAs).
Stammzelle	Stammzellen differenzieren sich im Laufe der Entwicklung zu Zellen mit spezifischen Funktionen (Knochen, Haut, Blutzellen, etc.). Von besonderem Interesse sind embryonale Stammzellen, da sie ‚pluripotent‘ angelegt sind und die Fähigkeit haben, sich in spezialisierte Zelltypen weiter zu entwickeln.
Toxin	Von einem Organismus synthetisiertes Molekül, das in einem anderen Organismus grundlegende biologische Vorgänge schädigt. Toxine sind meist Peptide oder Proteine, manche auch Alkaloide. Sie werden von Pilzen, Pflanzen und Tieren produziert. Bei Bakterien unterscheidet man Exo- und Endotoxine, komplex gebaute Verbindungen. Das erste entdeckte bakterielle Exotoxin ist das Diphtherietoxin.
Transformation	Übertragung genetischer Information eines ‚Donors‘ auf einen ‚Akzeptor‘. Die DNA des Donors kann auf verschiedene Weise in den Akzeptor (auch Wirtszelle genannt) eingebracht werden: Infektion mit einem natürlichen oder modifizierten Virus (auch als Transfektion bezeichnet); Transformation mit einem $\Rightarrow$ Plasmid (bei Bakterien); Mikro-Injektion der DNA in eine Zelle; ‚Beschuss‘ einer Wirtszelle mit Goldpartikeln, die mit der Donor-DNA beschichtet sind.
Transgene Organismen	Bezeichnet Organismen, die durch Einbringen einer ‚fremden‘ DNA genetisch verändert sind und dadurch potentiell nützliche Eigenschaften dazugewinnen.
Translokation	Bezeichnung für ‚großräumige‘ Mutationen, bei der ein oder mehrere Segmente eines Chromosoms auf ein ‚fremdes‘ Chromosom verschoben werden. Translokationen können dann schwerwiegende Krankheiten auslösen, wenn wichtige Gene dabei unter die Kontrolle eines falschen Regulators geraten. Translokationen spielen eine Rolle beim Burkitt-Lymphom, Philadelphia-Chromosom, Translokations-Trisomie.
Transkription	Die Umschreibung einer DNA-Sequenz in eine RNA-Sequenz. Diese Maßnahme hat mehrere wichtige Konsequenzen: 1. Erhaltung der ursprünglichen genetischen Information; 2. Herstellung vieler Kopien zum Gebrauch; 3. Kontrolle des An- oder Abschaltens eines Gens; 4. Kontrolle über die Lebensdauer eines Transkriptes.
Translation	Die Übersetzung der von einer messenger-RNA vorgegebenen Information in eine Peptidsequenz. Die Stätten der Proteinbiosynthese sind die Ribosomen. $\Rightarrow$ tRNAs, die jeweils für eine der 20 Aminosäuren spezifisch sind, übertragen diese Aminosäure auf die wachsende Peptidkette. Die Erkennung zwischen einem bestimmten Codon der messenger-RNA und der tRNA erfolgt durch Basenpaarung mit dem Anticodon (einem Basentriplett) in der Sequenz der tRNA) entsprechend den Regeln des $\Rightarrow$ Genetischen Codes.
Transposon	Genetisches Element, das mithilfe in seiner genetischen Information codierten Transposase an andere Stellen im Genom springen kann.
tRNA	Transferribonukleinsäuren dienen in der zellulären Proteinbiosynthese zur

	Ablesung der messenger-RNA. Die Codons eines Leserasters werden vom Anticodon der tRNA erkannt; diese überträgt die angekoppelten Aminosäuren entsprechend dem Genetischen Code auf die wachsende Peptidkette.
Viren	Viren bestehen aus einer Hülle (Proteinkapsel) und einer von ihr umschlossenen genetischen Information. Viren werden nach der Art ihrer genetischen Information eingeteilt. Sie kann aus DNA oder RNA bestehen und einzelsträngig oder doppelsträngig vorliegen. Viren infizieren eine geeignete Wirtszelle, in der ihre genetische Information dupliziert und zur Produktion weiterer Viruspartikel benutzt wird. (siehe auch $\Rightarrow$ Retroviren)
Zellzyklus	Jede Zelle durchläuft bei der Zellteilung einen geordneten Zyklus. In einer ersten Phase (G1) findet die Vorbereitung zur Reduplikation der DNA statt, die in der S-Phase abläuft. In der anschließenden G2-Phase wird die eigentliche Zellteilung vorbereitet, die dann in der Mitose (M-Phase) zur Aufteilung der Chromosomen, zur Teilung des Kernes und schließlich der Zelle selbst führt. Nach der Zellteilung kann sich in den beiden Tochterzellen die Zellteilung wiederholen oder sie können in eine Ruhephase (G0-Phase) eintreten. Eine Wiederholung der Zellteilung gibt es nicht für Nervenzellen oder Muskelzellen der gestreiften Muskulatur.

## **Anmerkungen und Quellenangaben**

Als Informationsquellen dienten ausschließlich die oben angeführten Zitate. Die angegebenen *web-linkes* sind auf dem Stand vom 12. Januar 2014. Die Auswahl überwiegend englischsprachiger Links aus Wikipedia berücksichtigt, dass dort breitere Informationen zu finden sind als in der deutschen Ausgabe.

Porträtfotos (Mendel, Avery, Watson, Crick, Arber, Berg, Sanger) wurden Wikimedia Commons (<http://commons.wikimedia.org/wiki/>) entnommen. ‚Freie‘ Vorlagen aus Wikimedia Commons wurden von mir modifiziert und eingearbeitet in folgende Abbildungen: Abb.1, Abb.2, Abb.6, Abb.15, Abb.19, Abb.26. Die geltenden Regelungen zur Benutzung sind erfüllt. Abb.17 und 18 waren Schenkungen eines Kollegen.